Yangon University of Economics

Post Graduate Diploma in Research Studies

(9th Batch)

DRS-212 Quantitative Research Methods

Second Quarter

Contents

Chapter 7	Confidence Intervals and Sample Size
Chapter 8	Hypothesis Testing
Chapter 9	Testing the Difference Between Two Means, Two Proportions, and Two Variances
Chapter 11	Other Chi-Square Tests
Chapter 12	Analysis of Variance

7

Confidence Intervals and Sample Size

© Fuse/Getty Images RF

STATISTICS TODAY

Stress and the College Student

A recent poll conducted by the mtvU/Associated Press found that 85% of college students reported that they experience stress daily. The study said, "It is clear that being stressed is a fact of life on college campuses today."

The study also reports that 74% of students' stress comes from school work, 71% from grades, and 62% from financial woes. The report stated that 2240 undergraduate students were selected and that the poll has a margin of error of ±3.0%.

In this chapter you will learn how to make a true estimate of a parameter, what is meant by the margin of error, and whether or not the sample size was large enough to represent all college students.

See Statistics Today—Revisited at the end of this chapter for more details.

OUTLINE

Introduction

7–1 Confidence Intervals for the Mean When σ Is Known

7–2 Confidence Intervals for the Mean When σ Is Unknown

7–3 Confidence Intervals and Sample Size for Proportions

7–4 Confidence Intervals for Variances and Standard Deviations

Summary

OBJECTIVES

After completing this chapter, you should be able to:

1. Find the confidence interval for the mean when σ is known.
2. Determine the minimum sample size for finding a confidence interval for the mean.
3. Find the confidence interval for the mean when σ is unknown.
4. Find the confidence interval for a proportion.
5. Determine the minimum sample size for finding a confidence interval for a proportion.
6. Find a confidence interval for a variance and a standard deviation.

Introduction

One aspect of inferential statistics is **estimation,** which is the process of estimating the value of a parameter from information obtained from a sample. For example, consider the following statements:

> "For each dollar you pay in county property tax, 22 cents covers the cost of incarcerating prisoners." (*Pittsburgh City Paper*)
>
> "The average amount employees and employers pay for health insurance is $11,664 per year." *(USA TODAY)*
>
> "For people who were asked if they won $5,000 tomorrow, 54% of them said that they would use it to pay off their debts." (ING U.S. Survey)
>
> "The average amount spent by a TV Super Bowl viewer is $63.87." (Retail Advertising and Marketing Association)
>
> "Eight percent of the people surveyed in the United States said that they participate in skiing in the winter time." (IMRE sports)
>
> "Consumers spent an average of $126 for Valentine's Day this year." (Nationa1 Retail Federation)

Since the populations from which these values were obtained are large, these values are only *estimates* of the true parameters and are derived from data collected from samples.

The statistical procedures for estimating the population mean, proportion, variance, and standard deviation will be explained in this chapter.

An important question in estimation is that of sample size. How large should the sample be in order to make an accurate estimate? This question is not easy to answer since the size of the sample depends on several factors, such as the accuracy desired and the probability of making a correct estimate. The question of sample size will be explained in this chapter also.

Inferential statistical techniques have various **assumptions** that must be met before valid conclusions can be obtained. One common assumption is that the samples must be randomly selected. Chapter 1 explains how to obtain a random sample. The other common assumption is that either the sample size must be greater than or equal to 30 or the population must be normally or approximately normally distributed if the sample size is less than 30.

To check this assumption, you can use the methods explained in Chapter 6. Just for review, the methods are to check the histogram to see if it is approximately bell-shaped, check for outliers, and if possible, generate a normal quantile plot and see whether the points fall close to a straight line. (*Note:* An area of statistics called nonparametric statistics does not require the variable to be normally distributed.)

7–1 Confidence Intervals for the Mean When σ Is Known

OBJECTIVE ①

Find the confidence interval for the mean when σ is known.

The main objective of this section is to show the procedure of estimating the value of an unknown population mean when the standard deviation of the population is known.

Suppose a college president wishes to estimate the average age of students attending classes this semester. The president could select a random sample of 100 students and find the average age of these students, say, 22.3 years. From the sample mean, the president could infer that the average age of all the students is 22.3 years. This type of estimate is called a *point estimate.*

A **point estimate** is a specific numerical value estimate of a parameter. The best point estimate of the population mean μ is the sample mean $\overline{X}$.

You might ask why other measures of central tendency, such as the median and mode, are not used to estimate the population mean. The reason is that the means of samples vary less than other statistics (such as medians and modes) when many samples are selected from the same population. Therefore, the sample mean is the best estimate of the population mean.

Sample measures (i.e., statistics) are used to estimate population measures (i.e., parameters). These statistics are called **estimators.** As previously stated, the sample mean is a better estimator of the population mean than the sample median or sample mode.

A good estimator should satisfy the three properties described next.

Three Properties of a Good Estimator

1. The estimator should be an **unbiased estimator.** That is, the expected value or the mean of the estimates obtained from samples of a given size is equal to the parameter being estimated.
2. The estimator should be consistent. For a **consistent estimator,** as sample size increases, the value of the estimator approaches the value of the parameter estimated.
3. The estimator should be a **relatively efficient estimator.** That is, of all the statistics that can be used to estimate a parameter, the relatively efficient estimator has the smallest variance.

Confidence Intervals

As stated in Chapter 6, the sample mean will be, for the most part, somewhat different from the population mean due to sampling error. Therefore, you might ask a second question: How good is a point estimate? The answer is that there is no way of knowing how close a particular point estimate is to the population mean.

This answer creates some doubt about the accuracy of point estimates. For this reason, statisticians prefer another type of estimate, called an *interval estimate.*

An **interval estimate** of a parameter is an interval or a range of values used to estimate the parameter. This estimate may or may not contain the value of the parameter being estimated.

In an interval estimate, the parameter is specified as being between two values. For example, an interval estimate for the average age of all students might be $21.9 < \mu < 22.7$, or 22.3 ± 0.4 years.

Either the interval contains the parameter or it does not. A degree of confidence (usually a percent) must be assigned before an interval estimate is made. For instance, you may wish to be 95% confident that the interval contains the true population mean. Another question then arises. Why 95%? Why not 99 or 99.5%?

If you desire to be more confident, such as 99 or 99.5% confident, then you must make the interval larger. For example, a 99% confidence interval for the mean age of college students might be $21.7 < \mu < 22.9$, or 22.3 ± 0.6. Hence, a tradeoff occurs. To be more confident that the interval contains the true population mean, you must make the interval wider.

Historical Notes

Point and interval estimates were known as long ago as the late 1700s. However, it wasn't until 1937 that a mathematician, J. Neyman, formulated practical applications for them.

The **confidence level** of an interval estimate of a parameter is the probability that the interval estimate will contain the parameter, assuming that a large number of samples are selected and that the estimation process on the same parameter is repeated.

A **confidence interval** is a specific interval estimate of a parameter determined by using data obtained from a sample and by using the specific confidence level of the estimate.

Intervals constructed in this way are called *confidence intervals.* Three common confidence intervals are used: the 90%, the 95%, and the 99% confidence intervals.

The algebraic derivation of the formula for determining a confidence interval for a mean will be shown later. A brief intuitive explanation will be given first.

The central limit theorem states that when the sample size is large, approximately 95% of the sample means of same-size samples taken from a population will fall within ±1.96 standard errors of the population mean, that is,

$$\mu \pm 1.96\left(\frac{\sigma}{\sqrt{n}}\right)$$

Now, if a specific sample mean is selected, say, $\overline{X}$, there is a 95% probability that the interval $\mu \pm 1.96(\sigma/\sqrt{n})$ contains $\overline{X}$. Likewise, there is a 95% probability that the interval specified by

$$\overline{X} \pm 1.96\left(\frac{\sigma}{\sqrt{n}}\right)$$

will contain μ, as will be shown later. Stated another way,

$$\overline{X} - 1.96\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \overline{X} + 1.96\left(\frac{\sigma}{\sqrt{n}}\right)$$

Hence, you can be 95% confident that the population mean is contained within that interval when the values of the variable are normally distributed in the population.

Interesting Fact

A postal worker who delivers mail walks on average 5.2 miles per day.

The value used for the 95% confidence interval, 1.96, is obtained from Table E in Appendix A. For a 99% confidence interval, the value 2.58 is used instead of 1.96 in the formula. This value is also obtained from Table E and is based on the standard normal distribution. Since other confidence intervals are used in statistics, the symbol $z_{\alpha/2}$ (read "zee sub alpha over two") is used in the general formula for confidence intervals. The Greek letter α (alpha) represents the total area in both tails of the standard normal distribution curve, and $\alpha/2$ represents the area in each one of the tails. The value $z_{\alpha/2}$ is called a *critical value*.

The relationship between α and the confidence level is that the stated confidence level is the percentage equivalent to the decimal value of $1 - \alpha$, and vice versa. When the 95% confidence interval is to be found, $\alpha = 0.05$, since $1 - 0.05 = 0.95$, or 95%. When $\alpha = 0.01$, then $1 - \alpha = 1 - 0.01 = 0.99$, and the 99% confidence interval is being calculated.

Formula for the Confidence Interval of the Mean for a Specific α When σ Is Known

$$\overline{X} - z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \overline{X} + z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

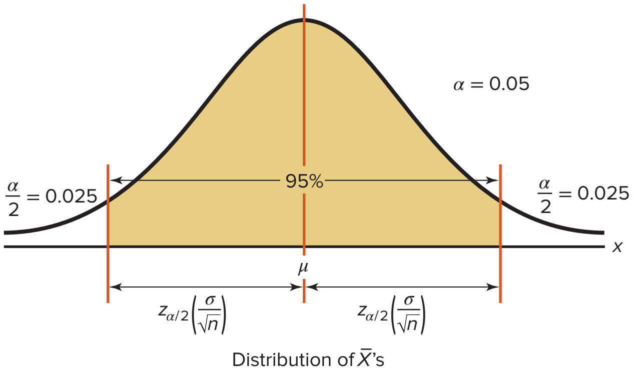
For a 90% confidence interval, $z_{\alpha/2} = 1.65$; for a 95% confidence interval, $z_{\alpha/2} = 1.96$; and for a 99% confidence interval, $z_{\alpha/2} = 2.58$.

The term $z_{\alpha/2}(\sigma/\sqrt{n})$ is called the *margin of error* (also called the *maximum error of the estimate*). For a specific value, say, $\alpha = 0.05$, 95% of the sample means will fall within this error value on either side of the population mean, as previously explained. See Figure 7–1.

The **margin of error**, also called the *maximum error of the estimate,* is the maximum likely difference between the point estimate of a parameter and the actual value of the parameter.

A more detailed explanation of the margin of error follows Examples 7–1 and 7–2, which illustrate the computation of confidence intervals.

FIGURE 7–1
95% Confidence Interval



$\alpha = 0.05$
$\frac{\alpha}{2} = 0.025$
95%
$\frac{\alpha}{2} = 0.025$
x
μ
$z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$
$z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$
Distribution of $\overline{X}$'s

Assumptions for Finding a Confidence Interval for a Mean When σ Is Known

1. The sample is a random sample.
2. Either $n \geq 30$ or the population is normally distributed when $n < 30$.

In this book, the assumptions will be stated in the exercises; however, when encountering statistics in other situations, you must check to see that these assumptions have been met before proceeding.

Some statistical techniques are called **robust.** This means that the distribution of the variable can depart somewhat from normality, and valid conclusions can still be obtained.

Rounding Rule for a Confidence Interval for a Mean When you are computing a confidence interval for a population mean by using *raw data,* round off to one more decimal place than the number of decimal places in the original data. When you are computing a confidence interval for a population mean by using a sample mean and a standard deviation, round off to the same number of decimal places as given for the mean.

EXAMPLE 7–1 Days It Takes to Sell an Aveo

A researcher wishes to estimate the number of days it takes an automobile dealer to sell a Chevrolet Aveo. A random sample of 50 cars had a mean time on the dealer's lot of 54 days. Assume the population standard deviation to be 6.0 days. Find the best point estimate of the population mean and the 95% confidence interval of the population mean.

Source: Based on information obtained from Power Information Network.

SOLUTION

The best point estimate of the population mean is 54 days. For the 95% confidence interval use $z = 1.96$.

$$\overline{X} - z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \overline{X} + z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$54 - 1.96\left(\frac{6.0}{\sqrt{50}}\right) < \mu < 54 + 1.96\left(\frac{6.0}{\sqrt{50}}\right)$$

$$54 - 1.7 < \mu < 54 + 1.7$$

$$52.3 < \mu < 55.7 \text{ or } 54 \pm 1.7$$

Hence, one can say with 95% confidence that the interval between 52.3 and 55.7 days does contain the population mean, based on a sample of 50 automobiles.

EXAMPLE 7–2 Number of Patients

A large urgent care center with 4 doctors found that they can see an average of 18 patients per hour. Assume the standard deviation is 3.2. A random sample of 42 hours was selected. Find the 99% confidence interval of the mean.

SOLUTION

The best point estimate of the population mean is 18. The 99% confidence level is

$$\overline{X} - z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \overline{X} + z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$18 - 2.58\left(\frac{3.2}{\sqrt{42}}\right) < \mu < 18 + 2.58\left(\frac{3.2}{\sqrt{42}}\right)$$

$$18 - 1.3 < \mu < 18 + 1.3$$

$$16.7 < \mu < 19.3$$

$$17 < \mu < 19 \text{ (rounded)}$$

Hence, one can be 99% confident (rounded values) that the mean number of patients that the center can care for in 1 hour is between 17 and 19.

Another way of looking at a confidence interval is shown in Figure 7–2. According to the central limit theorem, approximately 95% of the sample means fall within 1.96 standard deviations of the population mean if the sample size is 30 or more, or if σ is known when n is less than 30 and the population is normally distributed. If it were possible to build a confidence interval about each sample mean, as was done in Examples 7–1 and 7–2 for μ, then 95% of these intervals would contain the population mean, as shown in Figure 7–3. Hence, you can be 95% confident that an interval built around a specific sample mean would contain the population mean. If you desire to be 99% confident, you must enlarge the confidence intervals so that 99 out of every 100 intervals contain the population mean.

Since other confidence intervals (besides 90, 95, and 99%) are sometimes used in statistics, an explanation of how to find the values for $z_{\alpha/2}$ is necessary. As stated previously, the Greek letter α represents the total of the areas in both tails of the normal distribution. The value for α is found by subtracting the decimal equivalent for the desired confidence level from 1. For example, if you wanted to find the 98% confidence interval, you would change 98% to 0.98 and find $\alpha = 1 - 0.98$, or 0.02. Then $\alpha/2$ is obtained by dividing α by 2. So $\alpha/2$ is 0.02/2, or 0.01. Finally, $z_{0.01}$ is the z value that will give an area of 0.01 in the right tail of the standard normal distribution curve. See Figure 7–4.

Once $\alpha/2$ is determined, the corresponding $z_{\alpha/2}$ value can be found by using the procedure shown in Chapter 6, which is reviewed here. To get the $z_{\alpha/2}$ value for a 98%

FIGURE 7–2
95% Interval for Sample Means

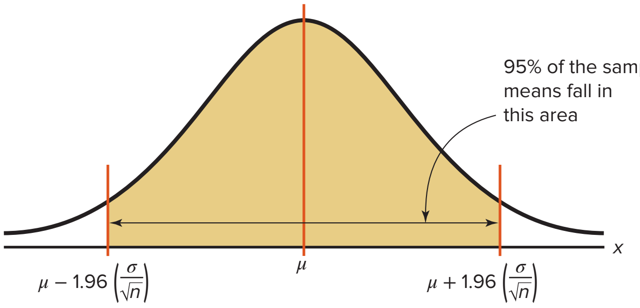


FIGURE 7–3
95% Confidence Intervals for Each Sample Mean

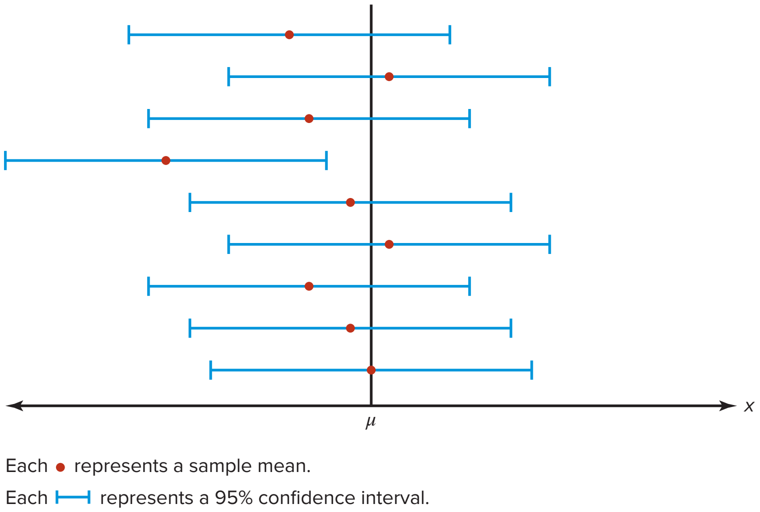


FIGURE 7–4
Finding $\alpha/2$ for a 98% Confidence Interval

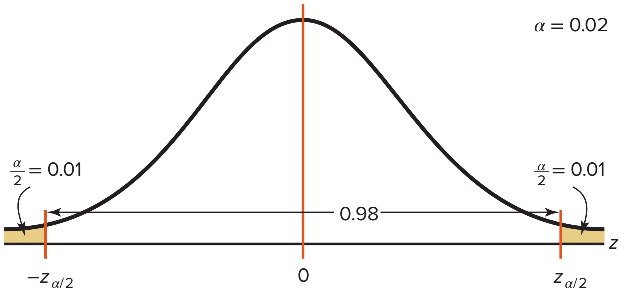
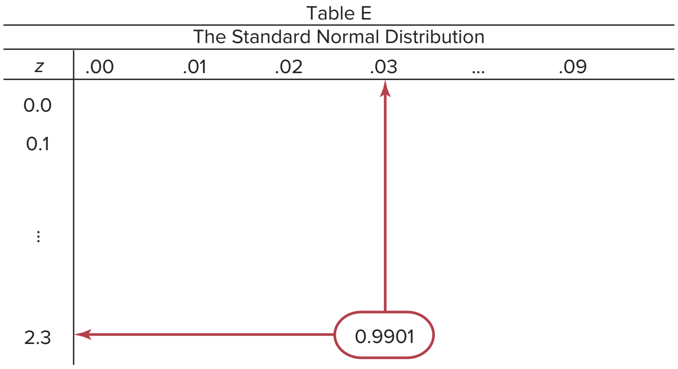


FIGURE 7–5
Finding $z_{\alpha/2}$ for a 98% Confidence Interval



confidence interval, subtract 0.01 from 1.0000 to get 0.9900. Next, locate the area that is closest to 0.9900 (in this case, 0.9901) in Table E, and then find the corresponding z value. In this example, it is 2.33. See Figure 7–5.

For confidence intervals, only the positive z value is used in the formula.

When the original variable is normally distributed and σ is known, the standard normal distribution can be used to find confidence intervals regardless of the size of the sample. When $n \geq 30$, the distribution of means will be approximately normal even if the original distribution of the variable departs from normality.

When σ is unknown, s can be used as an estimate of σ, but a different distribution is used for the critical values. This method is explained in Section 7–2.

EXAMPLE 7–3 Credit Union Assets

The following data represent a random sample of the assets (in millions of dollars) of 30 credit unions in southwestern Pennsylvania. Assume the population standard deviation is 14.405. Find the 90% confidence interval of the mean.

12.23	16.56	4.39
2.89	1.24	2.17
13.19	9.16	1.42
73.25	1.91	14.64
11.59	6.69	1.06
8.74	3.17	18.13
7.92	4.78	16.85
40.22	2.42	21.58
5.01	1.47	12.24
2.27	12.77	2.76

Source: Pittsburgh Post Gazette.

SOLUTION

Step 1 Find the mean for the data. Use the formula shown in Chapter 3 or your calculator. The mean $\overline{X} = 11.091$. Assume the standard deviation of the population is 14.405.

Step 2 Find $\alpha/2$. Since the 90% confidence interval is to be used, $\alpha = 1 - 0.90 = 0.10$, and

$$\frac{\alpha}{2} = \frac{0.10}{2} = 0.05$$

Step 3 Find $z_{\alpha/2}$. Subtract 0.05 from 1.000 to get 0.9500. The corresponding z value obtained from Table E is 1.65. (*Note:* This value is found by using the z value for an area between 0.9495 and 0.9505. A more precise z value obtained mathematically is 1.645 and is sometimes used; however, 1.65 will be used in this text.)

Step 4 Substitute in the formula

$$\overline{X} - z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \overline{X} + z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$11.091 - 1.65\left(\frac{14.405}{\sqrt{30}}\right) < \mu < 11.091 + 1.65\left(\frac{14.405}{\sqrt{30}}\right)$$

$$11.091 - 4.339 < \mu < 11.091 + 4.339$$

$$6.752 < \mu < 15.430$$

Hence, one can be 90% confident that the population mean of the assets of all credit unions is between \$6.752 million and \$15.430 million, based on a sample of 30 credit unions.

Comment to Computer and Statistical Calculator Users

This chapter and subsequent chapters include examples using raw data. If you are using computer or calculator programs to find the solutions, the answers you get may vary somewhat from the ones given in the text. This is so because computers and calculators do not round the answers in the intermediate steps and can use 12 or more decimal places for computation. Also, they use more-exact critical values than those given in the tables in the back of this book.

When you are calculating other statistics, such as the z, t, χ^2, or F values (shown in this chapter and later chapters), it is permissible to carry out the values of means, variances, and standard deviations to more decimal places than specified by the rounding rules in Chapter 3. This will give answers that are closer to the calculator or computer values. These small discrepancies are part of statistics.

Sample Size

OBJECTIVE 2

Determine the minimum sample size for finding a confidence interval for the mean.

Sample size determination is closely related to statistical estimation. Quite often you ask, How large a sample is necessary to make an accurate estimate? The answer is not simple, since it depends on three things: the margin of error, the population standard deviation, and the degree of confidence. For example, how close to the true mean do you want to be (2 units, 5 units, etc.), and how confident do you wish to be (90, 95, 99%, etc.)? For the purposes of this chapter, it will be assumed that the population standard deviation of the variable is known or has been estimated from a previous study.

The formula for sample size is derived from the margin of error formula

$$E = z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

and this formula is solved for n as follows:

$$E\sqrt{n} = z_{\alpha/2}(\sigma)$$

$$\sqrt{n} = \frac{z_{\alpha/2}\cdot\sigma}{E}$$

Hence,

$$n = \left(\frac{z_{\alpha/2}\cdot\sigma}{E}\right)^2$$

Formula for the Minimum Sample Size Needed for an Interval Estimate of the Population Mean

$$n = \left(\frac{z_{\alpha/2}\cdot\sigma}{E}\right)^2$$

where E is the margin of error. If necessary, round the answer up to obtain a whole number. That is, if there is any fraction or decimal portion in the answer, use the next whole number for sample size n.

EXAMPLE 7–4 Automobile Thefts

A sociologist wishes to estimate the average number of automobile thefts in a large city per day within 2 automobiles. He wishes to be 99% confident, and from a previous study the standard deviation was found to be 4.2. How many days should he select to survey?

SOLUTION

Since $\alpha = 0.01$ (or $1 - 0.99$), $z_{\alpha/2} = 2.58$ and $E = 2$. Substitute in the formula

$$n = \left(\frac{z_{\alpha/2}\cdot\sigma}{E}\right)^2 = \left[\frac{(2.58)(4.2)}{2}\right]^2 = 29.35$$

Round the value up to 30. Therefore, to be 99% confident that the estimate is within 2 automobiles of the true mean, the sociologist needs to sample the thefts for at least 30 days.

In most cases in statistics, we round off; however, when determining sample size, we always round up to the next whole number.

Interesting Fact

It has been estimated that the amount of pizza consumed every day in the United States would cover a farm consisting of 75 acres.

Notice that when you are finding the sample size, the size of the population is irrelevant when the population is large or infinite or when sampling is done with replacement. In other cases, an adjustment is made in the formula for computing sample size. This adjustment is beyond the scope of this book.

The formula for determining sample size requires the use of the population standard deviation. What happens when σ is unknown? In this case, an attempt is made to estimate σ. One such way is to use the standard deviation s obtained from a sample taken previously as an estimate for σ. The standard deviation can also be estimated by dividing the range by 4.

Sometimes, interval estimates rather than point estimates are reported. For instance, you may read a statement: "On the basis of a sample of 200 families, the survey estimates that an American family of two spends an average of \$84 per week for groceries. One can be 95% confident that this estimate is accurate within \$3 of the true mean." This statement means that the 95% confidence interval of the true mean is

$$\$84 - \$3 < \mu < \$84 + \$3$$

$$\$81 < \mu < \$87$$

The algebraic derivation of the formula for a confidence interval is shown next. As explained in Chapter 6, the sampling distribution of the mean is approximately normal when large samples ($n \geq 30$) are taken from a population. Also,

$$z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

Furthermore, there is a probability of $1 - \alpha$ that a z will have a value between $-z_{\alpha/2}$ and $+z_{\alpha/2}$. Hence,

$$-z_{\alpha/2} < \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2}$$

By using algebra, the formula can be rewritten as

$$-z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} < \overline{X} - \mu < z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Subtracting $\overline{X}$ from both sides and from the middle gives

$$-\overline{X} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} < -\mu < -\overline{X} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Multiplying by -1 gives

$$\overline{X} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} > \mu > \overline{X} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Reversing the inequality yields the formula for the confidence interval:

$$\overline{X} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Applying the Concepts 7–1

How Many Tissues Should Be in a Box?

Assume you work for a corporation that makes facial tissues. The job you are presently working on requires you to decide how many tissues are to be put in the new automobile glove compartment boxes. Complete the following.

1. How will you decide on a reasonable number of tissues to put in the boxes?
2. When do people usually need tissues?

3. What type of data collection technique would you use?
4. Assume you found out that from your sample of 85 people, an average of about 57 tissues are used throughout the duration of a cold, with a population standard deviation of 15. Use a confidence interval to help you decide how many tissues will go in the boxes.
5. Explain how you decided how many tissues will go in the boxes.

See page 411 for the answers.

Exercises 7–1

1. What is the difference between a point estimate and an interval estimate of a parameter? Which is better? Why?

2. What information is necessary to calculate a confidence interval?

3. What is the margin of error?

4. What is meant by the 95% confidence interval of the mean?

5. What are three properties of a good estimator?

6. What statistic best estimates μ?

7. Find each.

a. $z_{\alpha/2}$ for the 99% confidence interval

b. $z_{\alpha/2}$ for the 98% confidence interval

c. $z_{\alpha/2}$ for the 95% confidence interval

d. $z_{\alpha/2}$ for the 90% confidence interval

e. $z_{\alpha/2}$ for the 94% confidence interval

8. What is necessary to determine the sample size?

9. Fuel Efficiency of Cars and Trucks Since 1975 the average fuel efficiency of U.S. cars and light trucks (SUVs) has increased from 13.5 to 25.8 mpg, an increase of over 90%! A random sample of 40 cars from a large community got a mean mileage of 28.1 mpg per vehicle. The population standard deviation is 4.7 mpg. Estimate the true mean gas mileage with 95% confidence.

Source: World Almanac 2012.

10. Fast-Food Bills for Drive-Thru Customers A random sample of 50 cars in the drive-thru of a popular fast food restaurant revealed an average bill of \$18.21 per car. The population standard deviation is \$5.92. Estimate the mean bill for all cars from the drive-thru with 98% confidence.

11. Overweight Men For a random sample of 60 overweight men, the mean of the number of pounds that they were overweight was 30. The standard deviation of the population is 4.2 pounds.

a. Find the best point estimate of the average number of excess pounds that they weighed.

b. Find the 95% confidence interval of the mean of these pounds.

c. Find the 99% confidence interval of these pounds.

d. Which interval is larger? Why?

12. Number of Jobs A sociologist found that in a random sample of 50 retired men, the average number of jobs they had during their lifetimes was 7.2. The population standard deviation is 2.1.

a. Find the best point estimate of the population mean.

b. Find the 95% confidence interval of the mean number of jobs.

c. Find the 99% confidence interval of the mean number of jobs.

d. Which is smaller? Explain why.

13. Number of Faculty The numbers of faculty at 32 randomly selected state-controlled colleges and universities with enrollment under 12,000 students are shown below. Use these data to estimate the mean number of faculty at all state-controlled colleges and universities with enrollment under 12,000 with 92% confidence. Assume $\sigma = 165.1$.

211	384	396	211	224	337	395	121	356
621	367	408	515	280	289	180	431	176
318	836	203	374	224	121	412	134	539
471	638	425	159	324				

Source: World Almanac.

14. Freshmen GPAs First-semester GPAs for a random selection of freshmen at a large university are shown. Estimate the true mean GPA of the freshman class with 99% confidence. Assume $\sigma = 0.62$.

1.9	3.2	2.0	2.9	2.7	3.3
2.8	3.0	3.8	2.7	2.0	1.9
2.5	2.7	2.8	3.2	3.0	3.8
3.1	2.7	3.5	3.8	3.9	2.7
2.0	2.8	1.9	4.0	2.2	2.8
2.1	2.4	3.0	3.4	2.9	2.1

15. Carbohydrate Grams in Commercial Subs The number of grams of carbohydrates in various commercially prepared 7-inch subs is recorded below. The population standard deviation is 6.46. Estimate the mean number of carbs in all similarly sized subs with 95% confidence.

63	67	61	64	51	42	56	70	61
55	60	55	57	60	60	66	55	58
70	65	49	51	61	54	50	55	56
53	65	68	63	48	54	56	57	

16. **Number of Farms** A random sample of the number of farms (in thousands) in various states follows. Estimate the mean number of farms per state with 90% confidence. Assume $\sigma = 31.0$.

47	95	54	33	64	4	8	57	9	80
8	90	3	49	4	44	79	80	48	16
68	7	15	21	52	6	78	109	40	50
29									

Source: New York Times Almanac.

17. **Gasoline Use** A random sample of 36 drivers used on average 749 gallons of gasoline per year. If the standard deviation of the population is 32 gallons, find the 95% confidence interval of the mean for all drivers. If a driver said that he used 803 gallons per year, would you believe that?

18. **Day Care Tuition** A random sample of 50 four-year-olds attending day care centers provided a yearly tuition average of $3987 and the population standard deviation of $630. Find the 90% confidence interval of the true mean. If a day care center were starting up and wanted to keep tuition low, what would be a reasonable amount to charge?

19. **Hospital Noise Levels** Noise levels at various area urban hospitals were measured in decibels. The mean of the noise levels in 84 randomly selected corridors was 61.2 decibels, and the standard deviation of the population was 7.9. Find the 95% confidence interval of the true mean.

Source: M. Bayo, A. Garcia, and A. Garcia, "Noise Levels in an Urban Hospital and Workers' Subjective Responses," *Archives of Environmental Health* 50(3): 247-51 (May-June 1995).

20. **Length of Growing Seasons** The growing seasons for a random sample of 35 U.S. cities were recorded, yielding a sample mean of 190.7 days and the population standard deviation of 54.2 days. Estimate for all U.S. cities the true mean of the growing season with 95% confidence.

Source: The Old Farmer's Almanac.

21. **Christmas Presents** How large a sample is needed to estimate the population mean for the amount of money a person spends on Christmas presents within $2 and be 95% confident? The standard deviation of the population is $7.50.

22. **Hospital Noise Levels** In the hospital study cited in Exercise 19, the mean noise level in 171 randomly selected ward areas was 58.0 decibels, and the population standard deviation was 4.8. Find the 90% confidence interval of the true mean.

Source: M. Bayo, A. Garcia, and A. Garcia, "Noise Levels in an Urban Hospital and Workers' Subjective Responses," *Archives of Environmental Health* 50(3): 247-51 (May-June 1995).

23. **Internet Viewing** A researcher wishes to estimate the average number of minutes per day a person spends on the Internet. How large a sample must she select if she wishes to be 90% confident that the population mean is within 10 minutes of the sample mean? Assume the population standard deviation is 42 minutes.

24. **Cost of Pizzas** A pizza shop owner wishes to find the 95% confidence interval of the true mean cost of a large cheese pizza. How large should the sample be if she wishes to be accurate to within $0.15? A previous study showed that the standard deviation of the price was $0.26.

25. **Water Temperature** If the variance of the water temperature in a lake is 28°, how many days should the researcher select to measure the temperature to estimate the true mean within 3° with 99% confidence?

26. **Undergraduate GPAs** It is desired to estimate the mean GPA of each undergraduate class at a large university. How large a sample is necessary to estimate the GPA within 0.25 at the 99% confidence level? The population standard deviation is 1.2.

Technology Step by Step

TI-84 Plus Step by Step

Finding a *z* Confidence Interval for the Mean (Data)

1. Enter the data into L_1.
2. Press **STAT** and move the cursor to TESTS.
3. Press **7** for ZInterval.
4. Move the cursor to Data and press **ENTER.**
5. Type in the appropriate values.
6. Move the cursor to Calculate and press **ENTER.**

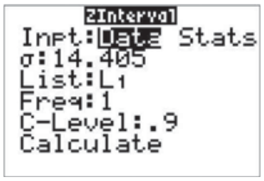
Example TI7–1

For Example 7–3 from the text, find the 90% confidence interval for the population mean, given the data values.

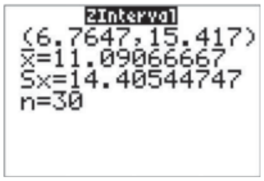
12.23	2.89	13.19	73.25	11.59	8.74	7.92	40.22	5.01	2.27
16.56	1.24	9.16	1.91	6.69	3.17	4.78	2.42	1.47	12.77
4.39	2.17	1.42	14.64	1.06	18.13	16.85	21.58	12.24	2.76

In the example, the population standard deviation σ is assumed to be 14.405. After the data values are entered in L_1 (step 1 above), enter 14.405 on the line for σ.

The 90% confidence interval is $6.765 < \mu < 15.417$. The difference between these limits and the ones in Example 7–3 is due to rounding.



```
ZInterval
Inpt:Data Stats
σ:14.405
List:L1
Freq:1
C-Level:.9
Calculate
```



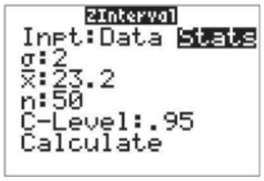
```
ZInterval
(6.7647,15.417)
x̄=11.09066667
Sx=14.40544747
n=30
```

Finding a *z* Confidence Interval for the Mean (Statistics)

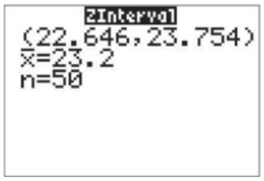
1. Press **STAT** and move the cursor to TESTS.
2. Press **7** for ZInterval.
3. Move the cursor to Stats and press **ENTER.**
4. Type in the appropriate values.
5. Move the cursor to Calculate and press **ENTER.**

Example TI7–2

Find the 95% confidence interval for the population mean, given $\sigma = 2$, $\overline{X} = 23.2$, and $n = 50$.



```
ZInterval
Inpt:Data Stats
σ:2
x̄:23.2
n:50
C-Level:.95
Calculate
```



```
ZInterval
(22.646,23.754)
x̄=23.2
n=50
```

The 95% confidence interval is $22.646 < \mu < 23.754$.

EXCEL
Step by Step

Finding a *z* Confidence Interval for the Mean

Excel has a procedure to compute the margin of error. But it does not compute confidence intervals. However, you may determine confidence intervals for the mean by using the MegaStat Add-in available in your online resources. If you have not installed this add-in, do so, following the instructions from the Chapter 1 Excel Step by Step.

Example XL7–1

Find the 95% confidence interval for the mean if $\sigma = 11$, using this sample:

43	52	18	20	25	45	43	21	42	32	24	32	19	25	26
44	42	41	41	53	22	25	23	21	27	33	36	47	19	20

1. Enter the data into an Excel worksheet.
2. Calculate the mean of your data using **=AVERAGE()**. Your result should be approximately 32.03.
3. From the toolbar, select Add-Ins, **MegaStat>Confidence Intervals/Sample Size**.
   Note: You may need to open **MegaStat** from the **MegaStat.xls** file on your computer's hard drive.
4. Enter the mean of the data, **32.03.**
5. Select *z* for the standard normal distribution.
6. Enter **11** for the standard deviation and **30** for *n*, the sample size.
7. Either type in or scroll to 95% for the Confidence Level, then click [OK].

The result of the procedure is shown next.

Confidence Interval—Mean

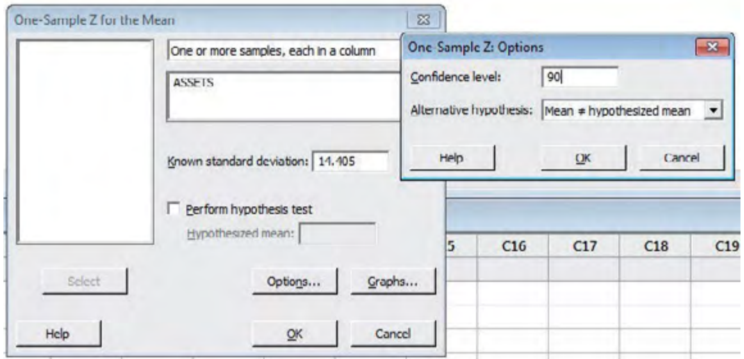
```
   95%   Confidence level
 32.03   Mean
    11   Standard deviation
    30   n
 1.960   z
 3.936   Half-width
35.966   Upper confidence limit
28.094   Lower confidence limit
```

MINITAB
Step by Step

Finding a *z* Confidence Interval for the Mean

For Example 7–3, find the 90% confidence interval estimate for the mean amount of assets for credit unions in southwestern Pennsylvania.

1. Maximize the worksheet, then enter the data into C1 of a MINITAB worksheet. Sigma is given as 14.405.
2. Select **Stat>Basic Statistics>1-Sample Z.**


One or more samples, each in a column

ASSETS

Known standard deviation: 14.405

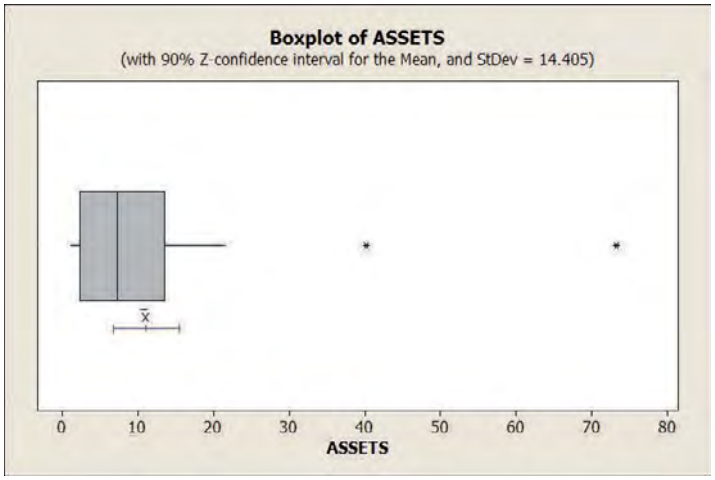
☐ Perform hypothesis test

Hypothesized mean:

Confidence level: 90

Alternative hypothesis: Mean ≠ hypothesized mean

   a) Select One or more samples, each in a column from the menu.
   b) Select C1 Assets by clicking the open space below this menu and double clicking the column name.
   c) Click on the box for Known standard deviation and type 14.405.
3. Click the [Options] button. In the dialog box make sure the Confidence Level is 90 and the alternative hypothesis is mean ≠ hypothesized mean. Click [OK].
4. Optional: Click [Graphs], then select Boxplot of data. The boxplot of these data will show possible outliers!
5. Click [OK] twice. The results will be displayed in the session window and boxplot.



One-Sample Z: Assets

```
The assumed standard deviation = 14.405

Variable   N    Mean   StDev   SE Mean       90% CI
Assets    30   11.09   14.41      2.63   (6.76, 15.42)
```

7–2 Confidence Intervals for the Mean When σ Is Unknown

OBJECTIVE 3

Find the confidence interval for the mean when σ is unknown.

When σ is known and the sample size is 30 or more, or the population is normally distributed if the sample size is less than 30, the confidence interval for the mean can be found by using the z distribution, as shown in Section 7–1. However, most of the time, the value of σ is not known, so it must be estimated by using s, namely, the standard deviation of the sample. When s is used, especially when the sample size is small, critical values greater than the values for $z_{\alpha/2}$ are used in confidence intervals in order to keep the interval at a given level, such as 95%. These values are taken from the *Student t distribution,* most often called the ***t* distribution.**

To use this method, the samples must be simple random samples, and the population from which the samples were taken must be normally or approximately normally distributed, or the sample size must be 30 or more.

Some important characteristics of the t distribution are described now.

Historical Notes

The *t* distribution was formulated in 1908 by an Irish brewing employee named W. S. Gosset. Gosset was involved in researching new methods of manufacturing ale. Because brewing employees were not allowed to publish results, Gosset published his finding using the pseudonym *Student;* hence, the *t* distribution is sometimes called *Student's t distribution.*

Characteristics of the *t* Distribution

The t distribution shares some characteristics of the standard normal distribution and differs from it in others. The t distribution is similar to the standard normal distribution in these ways:

1. It is bell-shaped.
2. It is symmetric about the mean.
3. The mean, median, and mode are equal to 0 and are located at the center of the distribution.
4. The curve approaches but never touches the x axis.

The t distribution differs from the standard normal distribution in the following ways:

1. The variance is greater than 1.
2. The t distribution is actually a family of curves based on the concept of *degrees of freedom*, which is related to sample size.
3. As the sample size increases, the t distribution approaches the standard normal distribution. See Figure 7–6.

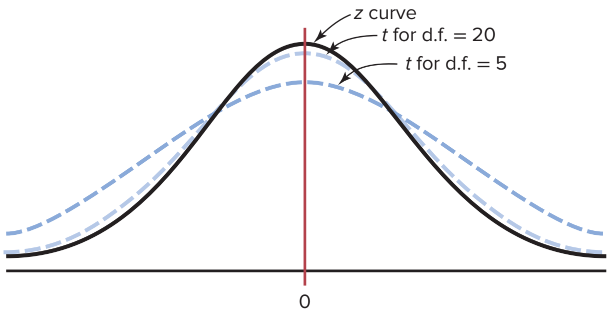
Many statistical distributions use the concept of degrees of freedom, and the formulas for finding the degrees of freedom vary for different statistical tests. The **degrees of freedom** are the number of values that are free to vary after a sample statistic has been computed, and they tell the researcher which specific curve to use when a distribution consists of a family of curves.

For example, if the mean of 5 values is 10, then 4 of the 5 values are free to vary. But once 4 values are selected, the fifth value must be a specific number to get a sum of 50, since $50 \div 5 = 10$. Hence, the degrees of freedom are $5 - 1 = 4$, and this value tells the researcher which t curve to use.

The symbol d.f. will be used for degrees of freedom. The degrees of freedom for a confidence interval for the mean are found by subtracting 1 from the sample size. That is, d.f. $= n - 1$. *Note:* For some statistical tests used later in this book, the degrees of freedom are not equal to $n - 1$.

FIGURE 7–6

The *t* Family of Curves



The formula for finding the confidence interval using the t distribution has a critical value $t_{\alpha/2}$.

The values for $t_{\alpha/2}$ are found in Table F in Appendix A. The top row of Table F, labeled Confidence Intervals, is used to get these values. The other two rows, labeled One tail and Two tails, will be explained in Chapter 8 and should not be used here.

Example 7–5 shows how to find the value in Table F for $t_{\alpha/2}$.

EXAMPLE 7–5

Find the $t_{\alpha/2}$ value for a 95% confidence interval when the sample size is 22.

SOLUTION

The d.f. = 22 − 1, or 21. Find 21 in the left column and 95% in the row labeled Confidence Intervals. The intersection where the two meet gives the value for $t_{\alpha/2}$, which is 2.080. See Figure 7–7.

FIGURE 7–7 Finding $t_{\alpha/2}$ for Example 7–5

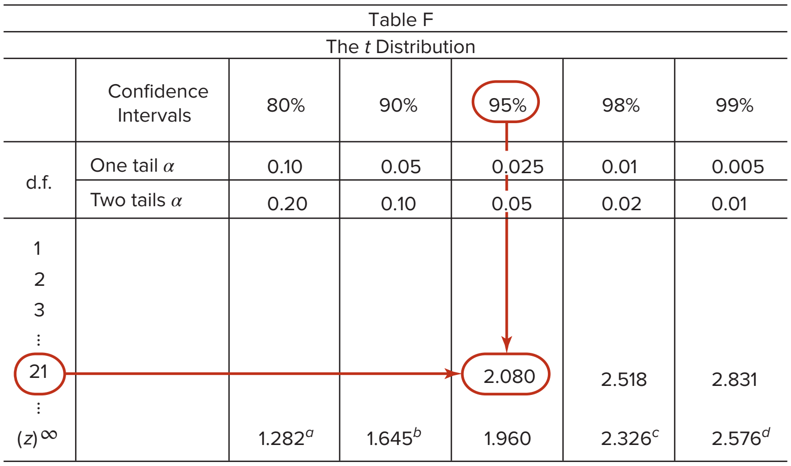


Table F

The *t* Distribution

d.f.	Confidence Intervals	80%	90%	95%	98%	99%
	One tail α	0.10	0.05	0.025	0.01	0.005
	Two tails α	0.20	0.10	0.05	0.02	0.01
1						
2						
3						
⋮						
21				2.080	2.518	2.831
⋮						
(z) ∞		1.282^a	1.645^b	1.960	2.326^c	2.576^d

When d.f. is greater than 30, it may fall between two table values. For example, if d.f. = 68, it falls between 65 and 70. Many texts say to use the closest value, for example, 68 is closer to 70 than 65; however, in this text a conservative approach is used. In this case, always round down to the nearest table value. In this case, 68 rounds down to 65.

Note: At the bottom of Table F where d.f. is large or ∞, the $z_{\alpha/2}$ values can be found for specific confidence intervals. The reason is that as the degrees of freedom increase, the t distribution approaches the standard normal distribution.

Examples 7–6 and 7–7 show how to find the confidence interval when you are using the t distribution.

The formula for finding a confidence interval for the mean by using the t distribution is given next.

Formula for a Specific Confidence Interval for the Mean When σ Is Unknown

$$\overline{X} - t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) < \mu < \overline{X} + t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

The degrees of freedom are $n - 1$.

The assumptions for finding a confidence interval for a mean when σ is unknown are given next.

Assumptions for Finding a Confidence Interval for a Mean When σ Is Unknown

1. The sample is a random sample.
2. Either $n \geq 30$ or the population is normally distributed when $n < 30$.

In this text, the assumptions will be stated in the exercises; however, when encountering statistics in other situations, you must check to see that these assumptions have been met before proceeding.

EXAMPLE 7–6 Temperature on Thanksgiving Day

A random sample of high temperatures for 12 recent Thanksgiving Days had an average of 42°F. Assume the variable is normally distributed and the standard deviation of the sample temperatures was 8°F. Find the 95% confidence interval of the population mean for the temperatures.

SOLUTION

$$\overline{X} = 42 \qquad s = 8 \qquad n = 12$$

Since σ is unknown and s must replace it, the t distribution (Table F) must be used for the confidence interval. Hence, with 11 degrees of freedom, $t_{\alpha/2} = 2.201$. The 95% confidence interval can be found by substituting in the formula:

$$\overline{X} - t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) < \mu < \overline{X} + t_{\alpha/2}\left(\frac{s}{n}\right)$$

$$42 - 2.201\left(\frac{8}{\sqrt{12}}\right) < \mu < 42 + 2.201\left(\frac{8}{\sqrt{12}}\right)$$

$$42 - 5.08 < \mu < 42 + 5.08$$

$$36.92 < \mu < 47.08$$

Therefore, one can be 95% confident that the population mean for the temperatures on Thanksgiving Days is between 36.92°F and 47.08°F.

EXAMPLE 7–7 Home Fires Started by Candles

The data represent a random sample of the number of home fires started by candles for the past several years. (Data are from the National Fire Protection Association.) Find the 99% confidence interval for the mean number of home fires started by candles each year.

5460 5900 6090 6310 7160 8440 9930

SOLUTION

Step 1 Find the mean and standard deviation for the data. Use the formulas in Chapter 3 or your calculator. The mean $\overline{X} = 7041.4$. The standard deviation $s = 1610.3$.

Step 2 Find $t_{\alpha/2}$ in Table F. Use the 99% confidence interval with d.f. = 6. It is 3.707.

Step 3 Substitute in the formula and solve.

$$\overline{X} - t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) < \mu < \overline{X} + t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

$$7041.4 - 3.707\left(\frac{1610.3}{\sqrt{7}}\right) < \mu < 7041.4 + 3.707\left(\frac{1610.3}{\sqrt{7}}\right)$$

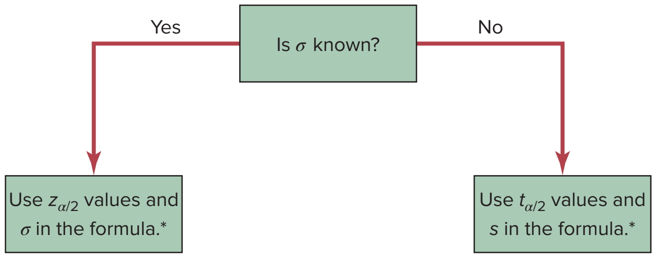
$$7041.4 - 2256.2 < \mu < 7041.4 + 2256.2$$

$$4785.2 < \mu < 9297.6$$

One can be 99% confident that the population mean number of home fires started by candles each year is between 4785.2 and 9297.6, based on a sample of home fires occurring over a period of 7 years.

FIGURE 7–8

When to Use the z or t Distribution


Is σ known?
Yes
No
Use $z_{\alpha/2}$ values and σ in the formula.*
Use $t_{\alpha/2}$ values and s in the formula.*

*If $n < 30$, the variable must be normally distributed.

Students sometimes have difficulty deciding whether to use $z_{\alpha/2}$ or $t_{\alpha/2}$ values when finding confidence intervals for the mean. As stated previously, when σ is known, $z_{\alpha/2}$ values can be used *no matter what the sample size is,* as long as the variable is normally distributed or $n \geq 30$. When σ is unknown and $n \geq 30$, then s can be used in the formula and $t_{\alpha/2}$ values can be used. Finally, when σ is unknown and $n < 30$, s is used in the formula and $t_{\alpha/2}$ values are used, as long as the variable is approximately normally distributed. These rules are summarized in Figure 7–8.

Applying the Concepts 7–2

Sport Drink Decision

Assume you get a new job as a coach for a sports team, and one of your first decisions is to choose the sports drink that the team will use during practices and games. You obtain a *Sports Report* magazine so you can use your statistical background to help you make the best decision. The following table lists the most popular sports drinks and some important information about each. Answer the following questions about the table.

Drink	Calories	Sodium	Potassium	Cost
Gatorade	60	110	25	$1.29
Powerade	68	77	32	1.19
All Sport	75	55	55	0.89
10-K	63	55	35	0.79
Exceed	69	50	44	1.59
1st Ade	58	58	25	1.09
Hydra Fuel	85	23	50	1.89

1. Would this be considered a small sample?
2. Compute the mean cost per container, and create a 90% confidence interval about that mean. Do all the costs per container fall inside the confidence interval? If not, which ones do not?
3. Are there any you would consider outliers?
4. How many degrees of freedom are there?
5. If cost is a major factor influencing your decision, would you consider cost per container or cost per serving?
6. List which drink you would recommend and why.

See page 411 for the answers.

Exercises 7–2

1. What are the properties of the *t* distribution?

2. What is meant by *degrees of freedom?*

3. Find the values for each.

a. $t_{\alpha/2}$ and $n = 18$ for the 99% confidence interval for the mean

b. $t_{\alpha/2}$ and $n = 23$ for the 95% confidence interval for the mean

c. $t_{\alpha/2}$ and $n = 15$ for the 98% confidence interval for the mean

d. $t_{\alpha/2}$ and $n = 10$ for the 90% confidence interval for the mean

e. $t_{\alpha/2}$ and $n = 20$ for the 95% confidence interval for the mean

4. When should the *t* distribution be used to find a confidence interval for the mean?

For Exercises 5 through 20, assume that all variables are approximately normally distributed.

5. Overtime Hours Worked A random sample of 15 registered nurses in a large hospital showed that they worked on average 44.2 hours per week. The standard deviation of the sample was 2.6. Estimate the mean of the population with 90% confidence.

6. Digital Camera Prices The prices (in dollars) for a particular model of digital camera with 18.0 megapixels and a f/3.5−5.6 zoom lens are shown here for 10 randomly selected online retailers. Estimate the true mean price for this particular model with 95% confidence.

999 1499 1997 398 591 498 798 849 449 348

7. Women Representatives in State Legislature A state representative wishes to estimate the mean number of women representatives per state legislature. A random sample of 17 states is selected, and the number of women representatives is shown. Based on the sample, what is the point estimate of the mean? Find the 90% confidence interval of the mean population. (*Note:* The population mean is actually 31.72, or about 32.) Compare this value to the point estimate and the confidence interval. There is something unusual about the data. Describe it and state how it would affect the confidence interval.

5 33 35 37 24
31 16 45 19 13
18 29 15 39 18
58 132

8. State Gasoline Taxes A random sample of state gasoline taxes (in cents) is shown here for 12 states. Use the data to estimate the true population mean gasoline tax with 90% confidence. Does your interval contain the national average of 44.7 cents?

38.4 40.9 67 32.5 51.5 43.4
38 43.4 50.7 35.4 39.3 41.4

Source: http://www.api.org/statistics/fueltaxes/

9. Calories in Candy Bars The number of calories per candy bar for a random sample of standard-size candy bars is shown below. Estimate the mean number of calories per candy bar with 98% confidence.

220 220 210 230 275
260 240 260 220 240
240 280 230 280

10. Dance Company Students The number of students who belong to the dance company at each of several randomly selected small universities is shown here. Estimate the true population mean size of a university dance company with 99% confidence.

21 25 32 22 28 30 29 30
47 26 35 26 35 26 28 28
32 27 40

11. Weights of Elephants A sample of 8 adult elephants had an average weight of 12,300 pounds. The standard deviation for the sample was 22 pounds. Find the 95% confidence interval of the population mean for the weights of adult elephants.

12. Thunderstorm Speeds A meteorologist who sampled 13 randomly selected thunderstorms found that the average speed at which they traveled across a certain state was 15.0 miles per hour. The standard deviation of the sample was 1.7 miles per hour. Find the 99% confidence interval of the mean. If a meteorologist wanted to use the highest speed to predict the times it would take storms to travel across the state in order to issue warnings, what figure would she likely use?

13. Work Time Lost due to Accidents At a large company, the Director of Research found that the average work time lost by employees due to accidents was 98 hours per year. She used a random sample of 18 employees. The standard deviation of the sample was 5.6 hours. Estimate the population mean for the number of hours lost due to accidents for the company, using a 95% confidence interval.

14. Social Networking Sites A recent survey of 8 randomly selected social networking sites has a mean of 13.1 million visitors for a specific month. The standard deviation is 4.1 million. Find the 95% confidence interval of the true mean.

Source: ComScore Media Matrix.

15. Stress Test Results For a group of 12 men subjected to a stress test situation, the average heart rate was 109 beats per minute. The standard deviation was 4. Find the 99% confidence interval of the population mean.

16. Hospital Noise Levels For a random sample of 24 operating rooms taken in the hospital study mentioned in Exercise 19 in Section 7–1, the mean noise level was 41.6 decibels, and the standard deviation was 7.5. Find the 95% confidence interval of the true mean of the noise levels in the operating rooms.

Source: M. Bayo, A. Garcia, and A. Garcia, "Noise Levels in an Urban Hospital and Workers' Subjective Responses," *Archives of Environmental Health* 50(3): 247-51 (May-June 1995).

17. Costs for a 30-Second Spot on Cable Television The approximate costs for 30-second randomly selected spots for various cable networks in a random selection of cities are shown. Estimate the true population mean cost for a 30-second advertisement on cable network with 90% confidence.

14 55 165 9 15 66 23 30 150
22 12 13 54 73 55 41 78

Source: www.spotrunner.com

18. Indy 500 Qualifier Speeds The speeds in miles per hour of eight randomly selected qualifiers for the Indianapolis 500 (in 2012) are listed below. Estimate the mean qualifying speed with 95% confidence.

224.037 226.484 222.891 222.929
223.422 225.172 226.240 223.684

19. NYSE Stock Prices An investing club randomly selects 15 NYSE stocks for consideration, and the prices per share are listed here. Estimate the mean price in dollars of all stocks with 95% confidence.

41.53 19.83 15.18 50.40 29.97
58.42 21.63 121.17 5.49 54.87
13.10 87.78 19.32 54.83 13.89

20. Unhealthy Days in Cities The number of unhealthy days based on the AQI (Air Quality Index) for a random sample of metropolitan areas is shown. Construct a 98% confidence interval based on the data.

61 12 6 40 27 38 93 5 13 40

Source: New York Times Almanac.

Extending the Concepts

21. Parking Meter Revenue A *one-sided confidence* interval can be found for a mean by using

$$\mu > \overline{X} - t_\alpha \frac{s}{\sqrt{n}} \qquad \text{or} \qquad \mu < \overline{X} + t_\alpha \frac{s}{\sqrt{n}}$$

where t_α is the value found under the row labeled One tail. Find two one-sided 95% confidence intervals of the population mean for the data shown, and interpret the answers. The data represent the daily revenues in dollars from 20 parking meters in a small municipality.

2.60 1.05 2.45 2.90
1.30 3.10 2.35 2.00
2.40 2.35 2.40 1.95
2.80 2.50 2.10 1.75
1.00 2.75 1.80 1.95

Technology Step by Step

TI-84 Plus
Step by Step

Finding a *t* Confidence Interval for the Mean (Data)

1. Enter the data into L$_1$.
2. Press **STAT** and move the cursor to TESTS.
3. Press **8** for TInterval.
4. Move the cursor to Data and press **ENTER.**
5. Type in the appropriate values.
6. Move the cursor to Calculate and press **ENTER.**

Finding a *t* Confidence Interval for the Mean (Statistics)

1. Press **STAT** and move the cursor to TESTS.
2. Press **8** for TInterval.
3. Move the cursor to Stats and press **ENTER**.
4. Type in the appropriate values.
5. Move the cursor to Calculate and press **ENTER**.

EXCEL
Step by Step

Finding a *t* Confidence Interval for the Mean

Excel has a procedure to compute the margin of error. But it does not compute confidence intervals. However, you may determine confidence intervals for the mean by using the MegaStat Add-in available in your online resources. If you have not installed this add-in, do so, following the instructions from the Chapter 1 Excel Step by Step.

Example XL7–2

Find the 95% confidence interval, using these sample data:

625 675 535 406 512 680 483 522 619 575

1. Enter the data into an Excel worksheet.
2. Calculate the mean of your data using **=AVERAGE().** Your result should be approximately 563.2.
3. Calculate the sample standard deviation using **=STDEV.S().** Your result should be approximately 87.9.
4. From the toolbar, select Add-Ins, **MegaStat>Confidence Intervals/Sample Size.** *Note:* You may need to open **MegaStat** from the **MegaStat.xls** file on your computer's hard drive.
5. Enter the mean of the data, **563.2.**
6. Select *t* for the *t* distribution.
7. Enter **87.9** for the standard deviation and **10** for *n*, the sample size.
8. Either type in or scroll to 95% for the Confidence Level, then click [OK].

The result of the procedure is shown next.

Confidence Interval—Mean

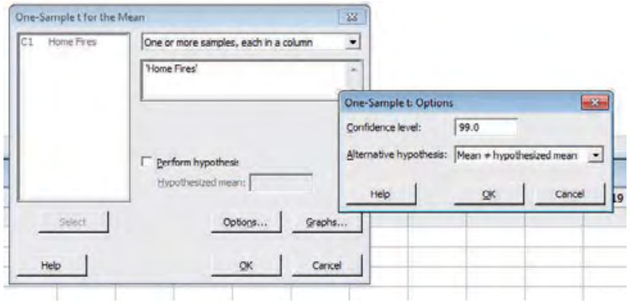
```
    95%    Confidence level
  563.2    Mean
   87.9    Standard deviation
     10    n
  2.262    t (d.f. = 9)
 62.880    Half-width
626.080    Upper confidence limit
500.320    Lower confidence limit
```

MINITAB
Step by Step

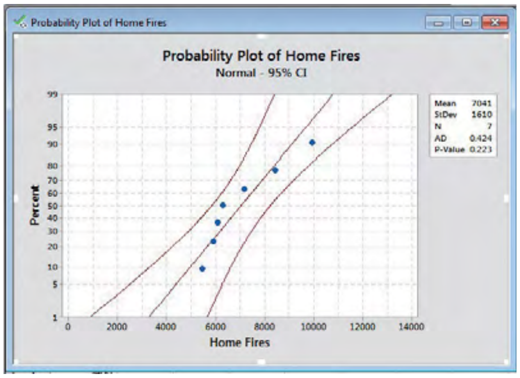
Find a *t* Interval for the Mean

For Example 7–7, find the 99% confidence interval for the mean number of home fires started by candles each year.

1. Type the data into C1 of a MINITAB worksheet. Name the column **Home Fires**.
2. Select One or more samples, each in a column from the menu and select C1 Home Fires by clicking the open space below this menu and double clicking the column name.
3. Click the [Options] button. In the dialog box, make sure the confidence level is 99 and the alternative hypothesis is mean ≠ hypothesized mean. Click [OK].
4. Click on [Options] and be sure the Confidence Level is 99 and the Alternative is not equal.
5. Click [OK] twice.



One or more samples, each in a column
'Home Fires'
☐ Perform hypothesis
Confidence level: 99.0
Alternative hypothesis: Mean ≠ hypothesized mean



6. Check for normality:
   a) Select **Graph>Probability Plot**, then Single.
   b) Select C1 Home Fires for the **Graph** variable. The normal plot is concave, a skewed distribution.
7. Click [OK].

In the session window you will see the results. The 99% confidence interval estimate for μ is between 4785 and 9298. The sample size, mean, standard deviation, and standard error of the mean are also shown.

However, this small sample appears to have a nonnormal population. The interval is less likely to contain the true mean.

```
One-Sample T: HOME FIRES
Variable      N     Mean     StDev     SE Mean       99% CI
HOME FIRES    7     7041     1610          609     (4785, 9298)
```

7–3 Confidence Intervals and Sample Size for Proportions

OBJECTIVE 4

Find the confidence interval for a proportion.

One of the most common types of confidence intervals is one that uses proportions. Many statistical studies involve finding a proportion of the population that has a certain characteristic. In this section, you will learn how to find the confidence interval for a population.

A *USA TODAY* Snapshots feature stated that 12% of the pleasure boats in the United States were named *Serenity*. The parameter 12% is called a **proportion.** It means that of all the pleasure boats in the United States, 12 out of every 100 are named *Serenity*. A proportion represents a part of a whole. It can be expressed as a fraction, decimal, or percentage. In this case, $12\% = 0.12 = \frac{12}{100}$ or $\frac{3}{25}$. Proportions can also represent probabilities. In this case, if a pleasure boat is selected at random, the probability that it is called *Serenity* is 0.12.

Proportions can be obtained from samples or populations. The following symbols will be used.

Symbols Used in Proportion Notation

$$p = \text{population proportion}$$

$$\hat{p} \text{ (read "}p\text{ hat")} = \text{sample proportion}$$

For a sample proportion,

$$\hat{p} = \frac{X}{n} \qquad \text{and} \qquad \hat{q} = \frac{n - X}{n} \qquad \text{or} \qquad \hat{q} = 1 - \hat{p}$$

where X = number of sample units that possess the characteristics of interest and n = sample size.

For example, in a study, 200 people were asked if they were satisfied with their jobs or professions; 162 said that they were. In this case, $n = 200$, $X = 162$, and $\hat{p} = X/n = 162/200 = 0.81$. It can be said that for this sample, 0.81, or 81%, of those surveyed were satisfied with their jobs or professions. The sample proportion is $\hat{p} = 0.81$.

The proportion of people who did not respond favorably when asked if they were satisfied with their jobs or professions constituted $\hat{q}$, where $\hat{q} = (n - X)/n$. For this survey, $\hat{q} = (200 - 162)/200 = 38/200$, or 0.19, or 19%.

When $\hat{p}$ and $\hat{q}$ are given in decimals or fractions, $\hat{p} + \hat{q} = 1$. When $\hat{p}$ and $\hat{q}$ are given in percentages, $\hat{p} + \hat{q} = 100\%$. It follows, then, that $\hat{q} = 1 - \hat{p}$, or $\hat{p} = 1 - \hat{q}$, when $\hat{p}$ and $\hat{q}$ are in decimal or fraction form. For the sample survey on job satisfaction, $\hat{q}$ can also be found by using $\hat{q} = 1 - \hat{p}$, or $1 - 0.81 = 0.19$.

Similar reasoning applies to population proportions; that is, $p = 1 - q$, $q = 1 - p$, and $p + q = 1$, when p and q are expressed in decimal or fraction form. When p and q are expressed as percentages, $p + q = 100\%$, $p = 100\% - q$, and $q = 100\% - p$.

EXAMPLE 7–8 Assault Victims

A random sample of 150 assault victims in a large city found that 45 took no self-protection actions against the criminal. Find $\hat{p}$ and $\hat{q}$, where $\hat{p}$ is the proportion of victims who took no self-protection action against the criminal.

SOLUTION

In this case, $X = 45$ and $n = 150$.

$$\hat{p} = \frac{X}{n} = \frac{45}{150} = 0.30 = 30\%$$

$$\hat{q} = \frac{n - X}{n} = \frac{150 - 45}{150} = \frac{105}{150} = 0.70 = 70\%$$

Also $\hat{q}$ can be found by using the formula $\hat{q} = 1 - \hat{p}$. In this case, $\hat{q} = 1 - 0.30 = 0.70$. So 70% of the victims used some self-defense action against the criminal.

As with means, the statistician, given the sample proportion, tries to estimate the population proportion. Point and interval estimates for a population proportion can be made by using the sample proportion. For a point estimate of p (the population proportion), $\hat{p}$ (the sample proportion) is used. On the basis of the three properties of a good estimator, $\hat{p}$ is unbiased, consistent, and relatively efficient. But as with means, one is not able to decide how good the point estimate of p is. Therefore, statisticians also use an interval estimate for a proportion, and they can assign a probability that the interval will contain the population proportion.

The confidence interval for a particular p is based on the sampling distribution of $\hat{p}$. When the sample size n is no more than 5% of the population size, the sampling distribution of $\hat{p}$ is approximately normal with a mean of p and a standard deviation of $\sqrt{pq/n}$, where $q = 1 - p$.

Confidence Intervals

To construct a confidence interval about a proportion, you must use the margin of error, which is

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Confidence intervals about proportions must meet the criteria that $n\hat{p} \geq 5$ and $n\hat{q} \geq 5$.

Formula for a Specific Confidence Interval for a Proportion

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

when $n\hat{p}$ and $n\hat{q}$ are each greater than or equal to 5.

Assumptions for Finding a Confidence Interval for a Population Proportion

1. The sample is a random sample.
2. The conditions for a binomial experiment are satisfied (see Chapter 5).

In this book, the assumptions will be stated in the exercises; however, when encountering statistics in other situations, you must check to see that these assumptions have been met before proceeding.

Rounding Rule for a Confidence Interval for a Proportion Round off to three decimal places.

EXAMPLE 7–9 Covering College Costs

A survey conducted by Sallie Mae and Gallup of 1404 respondents found that 323 students paid for their education by student loans. Find the 90% confidence interval of the true proportion of students who paid for their education by student loans.

SOLUTION

Step 1 Determine $\hat{p}$ and $\hat{q}$.

$$\hat{p} = \frac{X}{n} = \frac{323}{1404} = 0.23$$
$$\hat{q} = 1 - \hat{p} = 1.00 - 0.23 = 0.77$$

Step 2 Determine the critical value.

$$\alpha = 1 - 0.90 = 0.10$$
$$\frac{\alpha}{2} = \frac{0.10}{2} = 0.05$$
$$z_{\alpha/2} = 1.65$$

Step 3 Substitute in the formula

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$
$$0.23 - 1.65\sqrt{\frac{(0.23)(0.77)}{1404}} < p < 0.23 + 1.65\sqrt{\frac{(0.23)(0.77)}{1404}}$$
$$0.23 - 0.019 < p < 0.23 + 0.019$$
$$0.211 < p < 0.249$$

or
$$21.1\% < p < 24.9\%$$

Hence, you can be 90% confident that the percentage of students who pay for their college education by student loans is between 21.1 and 24.9%.

When a specific percentage is given, the percentage becomes $\hat{p}$ when it is changed to a decimal. For example, if the problem states that 12% of the applicants were men, then $\hat{p} = 0.12$.

EXAMPLE 7–10 Self-Driving Cars

A survey of 6000 adults found that 58% say that they would take a ride in a fully self-driving car. Find the 90% confidence interval of the true proportion who said that they would take a ride in a self-driving car.

Source: World Economic Forum and Boston Consulting Group.

SOLUTION

Step 1 Determine $\hat{p}$ and $\hat{q}$.

In this case, $\hat{p}$ is already given. It is 58%, or 0.58. Then $\hat{q}$ is $1 - \hat{p}$, or $1 - 0.58 = 0.42$.

Step 2 Determine the critical value.

$$\alpha = 1 - 0.90 = 0.10$$

$$\frac{\alpha}{2} = \frac{0.10}{2} = 0.05$$

$$z_{\alpha/2} = 1.65$$

Step 3 Substitute in the formula

$$\hat{p} = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$0.58 - 1.65\sqrt{\frac{(0.58)(0.42)}{6000}} < p < 0.58 + 1.65\sqrt{\frac{(0.58)(0.42)}{6000}}$$

$$0.58 - 0.0105 < p < 0.58 + 0.0105$$

$$0.570 < p < 0.591$$

$$57.0\,\% < p < 59.1\%$$

Hence, you can say with 90% confidence that the true percentage of people who would take a ride in a self-driving car is between 57.0% and 59.1%.

Sample Size for Proportions

To find the sample size needed to determine a confidence interval about a proportion, use this formula:

OBJECTIVE 5

Determine the minimum sample size for finding a confidence interval for a proportion.

Formula for Minimum Sample Size Needed for Interval Estimate of a Population Proportion

$$n = \hat{p}\hat{q}\left(\frac{z_{\alpha/2}}{E}\right)^2$$

If necessary, round up to obtain a whole number.

This formula can be found by solving the margin of error value for n in the formula

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

There are two situations to consider. First, if some approximation of $\hat{p}$ is known (e.g., from a previous study), that value can be used in the formula.

Second, if no approximation of $\hat{p}$ is known, you should use $\hat{p} = 0.5$. This value will give a sample size sufficiently large to guarantee an accurate prediction, given the confidence interval and the error of estimate. The reason is that when $\hat{p}$ and $\hat{q}$ are each 0.5, the product $\hat{p}\hat{q}$ is at maximum, as shown here.

$\hat{p}$	$\hat{q}$	$\hat{p}\hat{q}$
0.1	0.9	0.09
0.2	0.8	0.16
0.3	0.7	0.21
0.4	0.6	0.24
0.5	**0.5**	**0.25**
0.6	0.4	0.24
0.7	0.3	0.21
0.8	0.2	0.16
0.9	0.1	0.09

Using the maximum value yields the largest possible value of n for a given margin of error and for a given confidence interval.

The disadvantage of this method is that it can lead to a larger sample size than is necessary.

SPEAKING OF STATISTICS Does Success Bring Happiness?

W. C. Fields said, "Start every day off with a smile and get it over with."

Do you think people are happy because they are successful, or are they successful because they are happy people? A recent survey conducted by *Money* magazine showed that 34% of the people surveyed said that they were happy because they were successful; however, 63% said that they were successful because they were happy individuals. The people surveyed had an average household income of $75,000 or more. The margin of error was ±2.5%. Based on the information in this article, what would be the confidence interval for each percent?

© Pixtal/age fotostock RF

EXAMPLE 7–11 Land Line Phones

A researcher wishes to estimate, with 95% confidence, the proportion of people who did not have a land line phone. A previous study shows that 40% of those interviewed did not have a land line phone. The researcher wishes to be accurate within 2% of the true proportion. Find the minimum sample size necessary.

SOLUTION

Since $z_{\alpha/2} = 1.96$, $E = 0.02$, $\hat{p} = 0.40$, and $\hat{q} = 0.60$, then

$$n = \hat{p}\hat{q}\left(\frac{z_{\alpha/2}}{E}\right)^2 = (0.40)(0.60)\left(\frac{1.96}{0.02}\right)^2 = 2304.96$$

which, when rounded up, is 2305 people to interview. So the researcher must interview 2305 people.

EXAMPLE 7–12 Home Computers

In Example 7–11 assume that no previous study was done. Find the minimum sample size necessary to be accurate within 2% of the true population.

SOLUTION

Here we do not know the values of $\hat{p}$ and $\hat{q}$. So we use $\hat{p} = 0.5$ and $\hat{q} = 0.5$.

$$E = 0.02 \qquad \text{and} \qquad z_{\alpha/2} = 1.96$$

$$\begin{aligned} n &= \hat{p}\hat{q}\left(\frac{z_{\alpha/2}}{E}\right)^2 \\ &= (0.5)(0.5)\left(\frac{1.96}{0.02}\right)^2 \\ &= 2401 \end{aligned}$$

Hence, 2401 people must be interviewed when $\hat{p}$ is unknown. This is 96 more people than needed if $\hat{p}$ is known.

In determining the sample size, the size of the population is irrelevant. Only the degree of confidence and the margin of error are necessary to make the determination.

Applying the Concepts 7–3

Contracting Influenza

To answer the questions, use the following table describing the percentage of people who reported contracting influenza by gender and race/ethnicity.

Influenza		
Characteristic	**Percent**	**(95% CI)**
Gender		
Men	48.8	(47.1–50.5%)
Women	51.5	(50.2–52.8%)
Race/ethnicity		
Caucasian	52.2	(51.1–53.3%)
African American	33.1	(29.5–36.7%)
Hispanic	47.6	(40.9–54.3%)
Other	39.7	(30.8–48.5%)
Total	50.4	(49.3–51.5%)

Forty-nine states and the District of Columbia participated in the study. Weighted means were used. The sample size was 19,774. There were 12,774 women and 7000 men.

1. Explain what (95% CI) means.
2. How large is the margin of error for men reporting influenza?
3. What is the sample size?
4. How does the sample size affect the size of the confidence interval?
5. Would the confidence intervals be larger or smaller for a 90% CI, using the same data?
6. Where does the 51.5% influenza for women fit into its associated 95% CI?

See pages 411–412 for the answers.

Exercises 7–3

1. In each case, find $\hat{p}$ and $\hat{q}$.

a. $n = 80$ and $X = 40$

b. $n = 200$ and $X = 90$

c. $n = 130$ and $X = 60$

d. 25%

e. 42%

2. Find $\hat{p}$ and $\hat{q}$ for each situation.

a. $n = 60$ and $X = 35$

b. $n = 95$ and $X = 43$

c. 68%

d. 55%

e. 12%

3. Cyber Monday Shopping A survey of 1000 U.S. adults found that 33% of people said that they would get no work done on Cyber Monday since they would spend all day shopping online. Find the 95% confidence interval of the true proportion.

Source: hhgregg.

4. Manual Transmission Automobiles In 2014, six percent of the cars sold had a manual transmission. A random sample of college students who owned cars revealed the following: out of 122 cars, 26 had manual transmissions. Estimate the proportion of college students who drive cars with manual transmissions with 90% confidence.

5. Holiday Gifts A survey of 100 Americans found that 68% said they find it hard to buy holiday gifts that convey their true feelings. Find the 90% confidence interval of the population proportion.

Source: National Center for Education Statistics (www.nces.ed.gov).

6. Belief in Haunted Places A random sample of 205 college students was asked if they believed that places could be haunted, and 65 responded yes. Estimate the true proportion of college students who believe in the possibility of haunted places with 99% confidence.

According to *Time* magazine, 37% of all Americans believe that places can be haunted.

Source: Time magazine, Oct. 2006.

7. **Work Interruptions** Research by Steelcase found the average worker get interrupted every 11 minutes and takes 23 minutes to get back on task. From a random sample of 200 workers, 168 said they are interrupted every 11 minutes by email, texts, alerts, etc. Find the 90% confidence interval of the population proportion of workers who are interrupted every 11 minutes.

   Source: Entrepreneur.com

8. **Travel to Outer Space** A CBS News/*New York Times* poll found that 329 out of 763 randomly selected adults said they would travel to outer space in their lifetime, given the chance. Estimate the true proportion of adults who would like to travel to outer space with 92% confidence.

   Source: www.pollingreport.com

9. **High School Graduates Who Take the SAT** The national average for the percentage of high school graduates taking the SAT is 49%, but the state averages vary from a low of 4% to a high of 92%. A random sample of 300 graduating high school seniors was polled across a particular tristate area, and it was found that 195 had taken the SAT. Estimate the true proportion of high school graduates in this region who take the SAT with 95% confidence.

   Source: World Almanac.

10. **Educational Television** In a random sample of 200 people, 154 said that they watched educational television. Find the 90% confidence interval of the true proportion of people who watched educational television. If the television company wanted to publicize the proportion of viewers, do you think it should use the 90% confidence interval?

11. **Wi-Fi Access** A survey of 50 students in grades 4 through 12 found 68% have classroom Wi-Fi access. Find the 99% confidence interval of the population proportion.

12. **Students Who Major in Business** It has been reported that 20.4% of incoming freshmen indicate that they will major in business or a related field. A random sample of 400 incoming college freshmen was asked their preference, and 95 replied that they were considering business as a major. Estimate the true proportion of freshman business majors with 98% confidence. Does your interval contain 20.4?

    Source: New York Times Almanac.

13. **Smartphone Ownership** A recent survey of 349 people ages 18 to 29 found that 86% of them own a smartphone. Find the 99% confidence interval of the population proportion.

    Source: Pew Research Center.

14. **Home Internet Access** According to a study, 96% of adults ages 18–29 had internet access at home in 2015. A researcher wanted to estimate the proportion of undergraduate college students (18 to 29 years) with access, so she randomly sampled 180 undergraduates and found that 157 had access. Estimate the true proportion with 90% confidence.

15. **Overseas Travel** A researcher wishes to be 95% confident that her estimate of the true proportion of individuals who travel overseas is within 4% of the true proportion. Find the sample necessary if, in a prior study, a sample of 200 people showed that 40 traveled overseas last year. If no estimate of the sample proportion is available, how large should the sample be?

16. **Widows** A recent study indicated that 29% of the 100 women over age 55 in the study were widows.
    a. How large a sample must you take to be 90% confident that the estimate is within 0.05 of the true proportion of women over age 55 who are widows?
    b. If no estimate of the sample proportion is available, how large should the sample be?

17. **Direct Satellite Television** It is believed that 25% of U.S. homes have a direct satellite television receiver. How large a sample is necessary to estimate the true population of homes that do with 95% confidence and within 3 percentage points? How large a sample is necessary if nothing is known about the proportion?

    Source: New York Times Almanac.

18. **Obesity** Obesity is defined as a *body mass index* (BMI) of 30 kg/m^2 or more. A 95% confidence interval for the percentage of U.S. adults aged 20 years and over who were obese was found to be 22.4 to 23.5%. What was the sample size?

    Source: National Center for Health Statistics (www.cdc.gov/nchs).

19. **U.S. Fitness Guidelines** According to the World Almanac, 35.2% of men ages 18 to 44 meets the U.S. Fitness Guidelines. How large a sample is necessary to estimate the true proportion of men that are fit in this age group within 2 ½ percentage points with 90% confidence?

    Source: World Almanac 2016

20. **Diet Habits** A federal report indicated that 27% of children ages 2 to 5 years had a good diet—an increase over previous years. How large a sample is needed to estimate the true proportion of children with good diets within 2% with 95% confidence?

    Source: Federal Interagency Forum on Child and Family Statistics, *Washington Observer-Reporter.*

Extending the Concepts

21. Gun Control If a random sample of 600 people is selected and the researcher decides to have a margin of error of 4% on the specific proportion who favor gun control, find the degree of confidence. A recent study showed that 50% were in favor of some form of gun control.

22. Survey on Politics In a study, 68% of 1015 randomly selected adults said that they believe the Republicans favor the rich. If the margin of error was 3 percentage points, what was the confidence level used for the proportion?

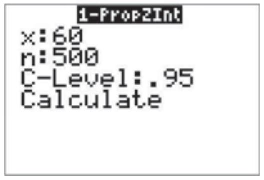
Source: USA TODAY.

Technology Step by Step

TI-84 Plus
Step by Step

Finding a Confidence Interval for a Proportion

1. Press **STAT** and move the cursor to TESTS.
2. Press **A (ALPHA, MATH)** for 1-PropZInt.
3. Type in the appropriate values.
4. Move the cursor to Calculate and press **ENTER.**

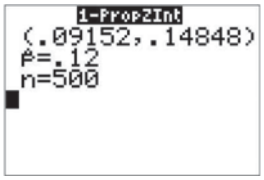


```
1-PropZInt
x:60
n:500
C-Level:.95
Calculate
```

Example TI7–3

Find the 95% confidence interval of p when $X = 60$ and $n = 500$.

The 95% confidence level for p is $0.09152 < p < 0.14848$. Also $\hat{p}$ is given.



```
1-PropZInt
(.09152,.14848)
p̂=.12
n=500
```

EXCEL
Step by Step

Finding a Confidence Interval for a Proportion

Excel has a procedure to compute the margin of error. But it does not compute confidence intervals. However, you may determine confidence intervals for a proportion by using the MegaStat Add-in available in your online resources. If you have not installed this add-in, do so, following the instructions from the Chapter 1 Excel Step by Step.

Example XL7–3

There were 500 nursing applications in a sample, including 60 from men. Find the 90% confidence interval for the true proportion of male applicants.

1. From the toolbar, select Add-Ins, **MegaStat>Confidence Intervals/Sample Size.** *Note:* You may need to open **MegaStat** from the **MegaStat.xls** file on your computer's hard drive.
2. In the dialog box, select Confidence interval—*p*.
3. Enter **60** in the box labeled p; p will automatically change to x.
4. Enter **500** in the box labeled n.
5. Either type in or scroll to 90% for the Confidence Level, then click [OK].

The result of the procedure is shown next.

Confidence Interval—Proportion

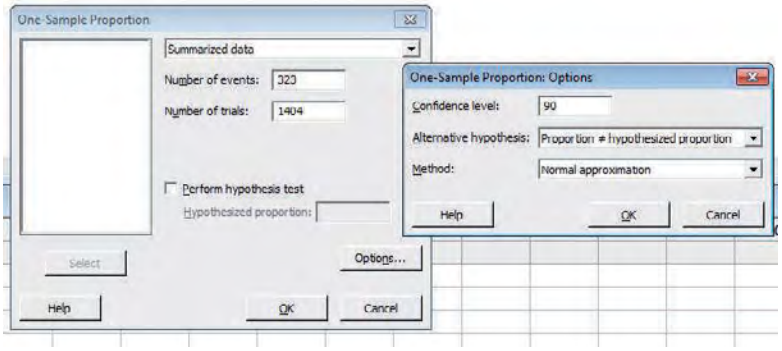
90%	Confidence level
0.12	Proportion
500	*n*
1.645	*z*
0.024	Half-width
0.144	Upper confidence limit
0.096	Lower confidence limit

MINITAB
Step by Step

Find a Confidence Interval for a Proportion

MINITAB will calculate a confidence interval, given the statistics from a sample *or* given the raw data. From Example 7–9 covering college costs, 323 out of 1404 respondents paid for their education by student loans. Find the 90% confidence interval of the true proportion of students who paid for their education by student loans.

1. Select **Stat>Basic Statistics>1 Proportion**.
2. In the drop down menu, select Summarized data. No data will be entered in the worksheet.
3. Click in the box for Number of trials and enter **1404.**
4. In the Number of events box, enter **323.**
5. Click on [Options].
6. Type **90** for the confidence level.



Summarized data
Number of events: 323
Number of trials: 1404
☐ Perform hypothesis test
Hypothesized proportion:
Confidence level: 90
Alternative hypothesis: Proportion ≠ hypothesized proportion
Method: Normal approximation

7. Select Proportion ≠ hypothesized proportion or the alternative hypothesis.
8. Select Normal Approximation for method.
9. Click [OK] twice.

The results for the confidence interval will be displayed in the session window.

Test and CI for One Proportion

```
Sample    X      N     Sample p          90% CI
1         323    1404  0.230057    (0.211582, 0.248532)
Using the normal approximation.
```

7–4 Confidence Intervals for Variances and Standard Deviations

OBJECTIVE 6

Find a confidence interval for a variance and a standard deviation.

In Sections 7–1 through 7–3 confidence intervals were calculated for means and proportions. This section will explain how to find confidence intervals for variances and standard deviations. In statistics, the variance and standard deviation of a variable are as important as the mean. For example, when products that fit together (such as pipes) are manufactured, it is important to keep the variations of the diameters of the products as small as possible; otherwise, they will not fit together properly and will have to be scrapped. In the manufacture of medicines, the variance and standard deviation of the medication in the pills play an important role in making sure patients receive the proper dosage. For these reasons, confidence intervals for variances and standard deviations are necessary.

To calculate these confidence intervals, a new statistical distribution is needed. It is called the **chi-square distribution.**

The chi-square variable is similar to the *t* variable in that its distribution is a family of curves based on the number of degrees of freedom. The symbol for chi-square is χ^2

© Royalty-Free/CORBIS

(Greek letter chi, pronounced "ki"). Several of the distributions are shown in Figure 7–9, along with the corresponding degrees of freedom. The chi-square distribution is obtained from the values of $(n-1)s^2/\sigma^2$ when random samples are selected from a normally distributed population whose variance is σ^2.

A chi-square variable cannot be negative, and the distributions are skewed to the right. At about 100 degrees of freedom, the chi-square distribution becomes somewhat symmetric. The area under each chi-square distribution is equal to 1.00, or 100%.

Historical Note

The χ^2 distribution with 2 degrees of freedom was formulated by a mathematician named Hershel in 1869 while he was studying the accuracy of shooting arrows at a target. Many other mathematicians have since contributed to its development.

A summary of the characteristics of the chi-square distribution is given next.

Characteristics of the Chi-Square Distribution

1. All chi-square values are greater than or equal to 0.
2. The chi-square distribution is a family of curves based on the degrees of freedom.
3. The area under each chi-square distribution curve is equal to 1.
4. The chi-square distributions are positively skewed.

FIGURE 7–9

The Chi-Square Family of Curves

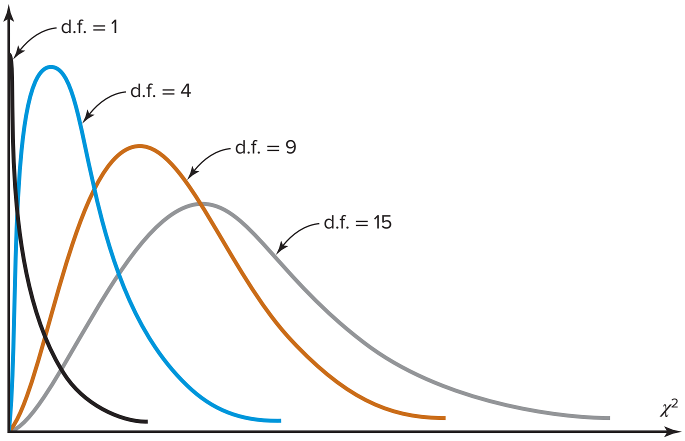


FIGURE 7–10

Chi-Square Distribution for d.f. $= n - 1$

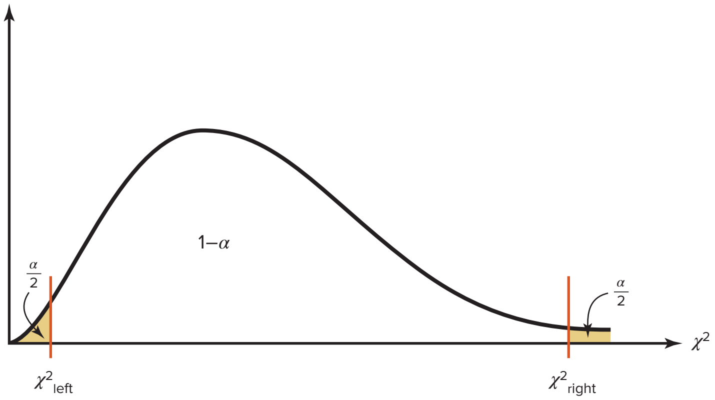


Table G in Appendix A gives the values for the chi-square distribution. These values are used in the denominators of the formulas for confidence intervals. Two different values are used in the formula because the distribution is not symmetric. One value is found on the left side of the table, and the other is on the right. See Figure 7–10.

Table G is set up similarly to the table used for the t distribution. The left column denotes the degrees of freedom, and the top row represents the area to the *right* of the critical value. For example, to find the table values corresponding to the 95% confidence interval, you must first change 95% to a decimal and subtract it from 1 ($1 - 0.95 = 0.05$). Then divide the answer by 2 ($\alpha/2 = 0.05/2 = 0.025$). This is the column on the right side of the table, used to get the values for χ^2_{right}. To get the value for χ^2_{left}, subtract the value of $\alpha/2$ from 1 ($1 - 0.05/2 = 0.975$). Finally, find the appropriate row corresponding to the degrees of freedom $n - 1$. A similar procedure is used to find the values for a 90 or 99% confidence interval.

EXAMPLE 7–13

Find the values for χ^2_{right} and χ^2_{left} for a 90% confidence interval when $n = 25$.

SOLUTION

To find χ^2_{right}, subtract $1 - 0.90 = 0.10$; then divide 0.10 by 2 to get 0.05.
To find χ^2_{left}, subtract $1 - 0.05 = 0.95$.
Then use the 0.95 and 0.05 columns with d.f. $= n - 1 = 25 - 1 = 24$. See Figure 7–11.

FIGURE 7–11 χ^2 Table for Example 7–13

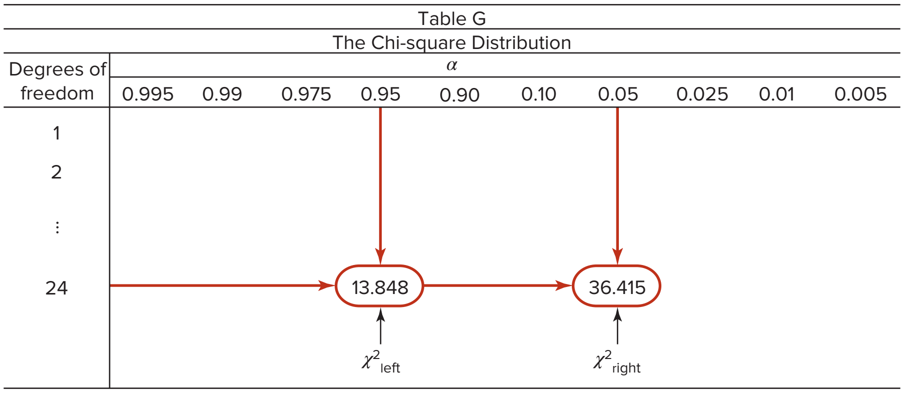


Table G

The Chi-square Distribution

Degrees of freedom	α									
	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1										
2										
⋮										
24				13.848			36.415			

χ^2_{left}
χ^2_{right}

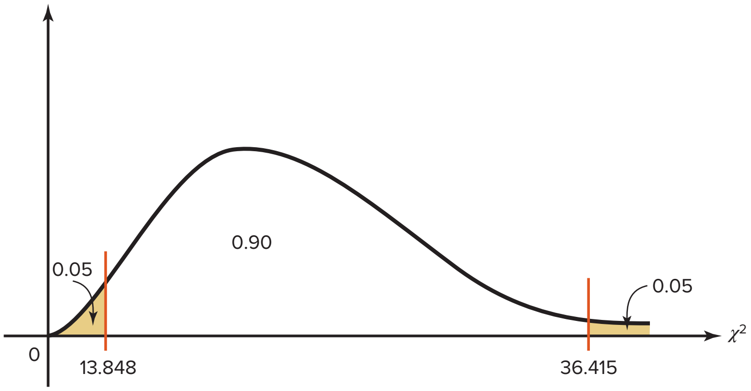
The values are

$$\chi^2_{\text{right}} = 36.415$$

$$\chi^2_{\text{left}} = 13.848$$

See Figure 7–12.

FIGURE 7–12 χ^2 Distribution for Example 7–13



If the number for the degrees of freedom is not given in the table, use the closest lower value in the table. For example, for d.f. = 53, use d.f. = 50. This is a conservative approach.

Useful estimates for σ^2 and σ are s^2 and s, respectively.

To find confidence intervals for variances and standard deviations, you must assume that the variable is normally distributed.

The formulas for the confidence intervals are shown here.

Formula for the Confidence Interval for a Variance

$$\frac{(n-1)s^2}{\chi^2_{\text{right}}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{\text{left}}}$$

$$\text{d.f.} = n - 1$$

Formula for the Confidence Interval for a Standard Deviation

$$\sqrt{\frac{(n-1)s^2}{\chi^2_{\text{right}}}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi^2_{\text{left}}}}$$

$$\text{d.f.} = n - 1$$

Recall that s^2 is the symbol for the sample variance and s is the symbol for the sample standard deviation. If the problem gives the sample standard deviation s, be sure to *square* it when you are using the formula. But if the problem gives the sample variance s^2, *do not square it* when you are using the formula, since the variance is already in square units.

Assumptions for Finding a Confidence Interval for a Variance or Standard Deviation

1. The sample is a random sample.
2. The population must be normally distributed.

In this text, the assumptions will be stated in the exercises; however, when encountering statistics in other situations, you must check to see that these assumptions have been met before proceeding.

Rounding Rule for a Confidence Interval for a Variance or Standard Deviation When you are computing a confidence interval for a population variance or standard deviation by using raw data, round off to one more decimal place than the number of decimal places in the original data.

When you are computing a confidence interval for a population variance or standard deviation by using a sample variance or standard deviation, round off to the same number of decimal places as given for the sample variance or standard deviation.

Example 7–14 shows how to find a confidence interval for a variance and standard deviation.

EXAMPLE 7–14 Women's Ages

A study of 30 women found the standard deviation of their ages was 5.2 years. Find the 95% confidence interval of the variance for the age variable. Assume the variable is normally distributed.

SOLUTION

Since $\alpha = 0.05$, the two critical values, respectfully, for the 0.025 and 0.975 levels for 29 degrees of freedom are 45.722 and 16.047.
Substitute in the formula:

$$\frac{(n-1)s^2}{\chi^2_{\text{right}}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{\text{left}}}$$

$$\frac{(30-1)(5.2)^2}{45.722} < \sigma^2 < \frac{(30-1)(5.2)^2}{16.047}$$

$$17.2 < \sigma^2 < 48.9$$

Hence, you can be 95% confident that the true variance for the ages is between 17.2 and 48.9 years.

For the standard deviation, the confidence interval is

$$\sqrt{17.2} < \sigma < \sqrt{48.9}$$

$$4.1 < \sigma < 7.0$$

Hence, you can be 95% confident that the true standard deviation of the age variable is between 4.1 and 7.0 based on a sample of 30 women.

EXAMPLE 7–15 Named Storms

Find the 90% confidence interval for the variance and standard deviation for the number of named storms per year in the Atlantic basin. A random sample of 10 years has been used. Assume the distribution is approximately normal.

10	5	12	11	13
15	19	18	14	16

Source: Atlantic Oceanographic and Meteorological Laboratory.

SOLUTION

Step 1 Find the variance for the data. Use the formulas in Chapter 3 or your calculator. The variance $s^2 = 16.9$.

Step 2 Find χ^2_{right} and χ^2_{left} from Table G in Appendix A, using $10 - 1 = 9$ degrees of freedom.

In this case, use $\alpha = 0.10$ and 0.95; $\chi^2_{\text{right}} = 3.325$; $\chi^2_{\text{left}} = 16.919$.

Step 3 Substitute in the formula.

$$\frac{(n-1)s^2}{\chi^2_{\text{right}}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{\text{left}}}$$

$$\frac{(10-1)(16.9)}{16.919} < \sigma^2 < \frac{(10-1)(16.9)}{3.325}$$

$$8.99 < \sigma^2 < 45.74$$

$$\sqrt{8.99} < \sigma < \sqrt{45.74}$$

$$3.0 < \sigma < 6.8$$

Hence, you can be 90% confident that the standard deviation for the number of named storms is between 3.0 and 6.8 based on a random sample of 10 years.

Note: If you are using the standard deviation instead (as in Example 7–14) of the variance, be sure to square the standard deviation when substituting in the formula.

Applying the Concepts 7–4

Ages of Presidents at the Time of Their Deaths

Shown are the ages (in years) of the Presidents at the times of their deaths.

67	90	83	85	73	80	78	79
68	71	53	65	74	64	77	56
66	63	70	49	57	71	67	71
58	60	72	67	57	60	90	63
88	78	46	64	81	93	93	

1. Do the data represent a population or a sample?
2. Select a random sample of 12 ages and find the variance and standard deviation.
3. Find the 95% confidence interval of the standard deviation.
4. Find the standard deviation of all the data values.
5. Does the confidence interval calculated in question 3 contain the standard deviation?
6. If it does not, give a reason why.
7. What assumption(s) must be considered for constructing the confidence interval in step 3?

See page 412 for the answers.

Exercises 7–4

1. What distribution must be used when computing confidence intervals for variances and standard deviations?

2. What assumption must be made when computing confidence intervals for variances and standard deviations?

3. Using Table G, find the values for χ^2_{left} and χ^2_{right}.

a. $\alpha = 0.05, n = 12$
b. $\alpha = 0.10, n = 20$
c. $\alpha = 0.05, n = 27$
d. $\alpha = 0.01, n = 6$
e. $\alpha = 0.10, n = 41$

4. Lifetimes of Wristwatches Find the 90% confidence interval for the variance and standard deviation for the lifetimes of inexpensive wristwatches if a random sample of 24 watches has a standard deviation of 4.8 months. Assume the variable is normally distributed. Do you feel that the lifetimes are relatively consistent?

5. Carbohydrates in Yogurt The number of carbohydrates (in grams) per 8-ounce serving of yogurt for each of a random selection of brands is listed below. Estimate the true population variance and standard deviation for the number of carbohydrates per 8-ounce serving of yogurt with 95% confidence. Assume the variable is normally distributed.

17 42 41 20 39 41 35 15 43
25 38 33 42 23 17 25 34

6. Carbon Monoxide Deaths A study of generation-related carbon monoxide deaths showed that a random sample of 6 recent years had a standard deviation of 4.1 deaths per year. Find the 99% confidence interval of the variance and standard deviation. Assume the variable is normally distributed.

Source: Based on information from Consumer Protection Safety Commission.

7. Pacemaker Batteries A manufacturer of pacemakers wants the standard deviation of the lifetimes of the batteries to be less than 1.8 months. A sample of 20 batteries had a standard deviation of 1.6 months. Assume the variable is normally distributed. Find the 95% confidence interval of the standard deviation of the batteries. Based on this answer, do you feel that 1.8 months is a reasonable estimate?

8. Age of College Students Find the 90% confidence interval for the variance and standard deviation of the ages of seniors at Oak Park College if a random sample of 24 students has a standard deviation of 2.3 years. Assume the variable is normally distributed.

9. Automobile Repairs The standard deviation of a sample of 15 automobile repairs at a local garage was $120.82. Assume the variable is normally distributed. Find the 90% confidence of the true variance and standard deviation.

10. Stock Prices A random sample of stock prices per share (in dollars) is shown. Find the 90% confidence interval for the variance and standard deviation for the prices. Assume the variable is normally distributed.

26.69	13.88	28.37	12.00
75.37	7.50	47.50	43.00
3.81	53.81	13.62	45.12
6.94	28.25	28.00	60.50
40.25	10.87	46.12	14.75

Source: Pittsburgh Tribune Review.

11. Cost of an Operation A medical researcher surveyed 11 hospitals and found that the standard deviation for the cost for removing a person's gall bladder was $53. Assume the variable is normally distributed. Based on this, find the 99% confidence interval of the population variance and standard deviation.

12. Home Ownership Rates The percentage rates of home ownership for 8 randomly selected states are listed below. Estimate the population variance and standard deviation for the percentage rate of home ownership with 99% confidence. Assume the variable is normally distributed.

66.0 75.8 70.9 73.9 63.4 68.5 73.3 65.9

Source: World Almanac.

13. Calories in a Standard Size Candy Bar Estimate the standard deviation in calories for these randomly selected standard-size candy bars with 95% confidence. (The number of calories is listed for each.) Assume the variable is normally distributed.

220 220 210 230 275 260 240
220 240 240 280 230 280 260

14. SAT Scores Estimate the variance in mean mathematics SAT scores by state, using the randomly selected scores listed below. Estimate with 99% confidence. Assume the variable is normally distributed.

490 502 211 209 499 565
469 543 572 550 515 500

Source: World Almanac 2012.

15. Daily Cholesterol Intake The American Heart Association recommends a daily cholesterol intake of less than 300 mg. Here are the cholesterol amounts in a random sample of single servings of grilled meats. Estimate the standard deviation in cholesterol with 95% confidence. Assume the variable is normally distributed.

90 200 80 105 95
85 70 105 115 110
100 225 125 130 145

Extending the Concepts

16. **Calculator Battery Lifetimes** A confidence interval for a standard deviation for large samples taken from a normally distributed population can be approximated by

$$s - z_{\alpha/2}\frac{s}{\sqrt{2n}} < \sigma < s + z_{\alpha/2}\frac{s}{\sqrt{2n}}$$

Find the 95% confidence interval for the population standard deviation of calculator batteries. A random sample of 200 calculator batteries has a standard deviation of 18 months.

Technology Step by Step

TI-84 Plus Step by Step

The TI-84 Plus does not have a built-in confidence interval for the variance or standard deviation. However, the downloadable program named SDINT is available in your online resources. Follow the instructions online for downloading the program.

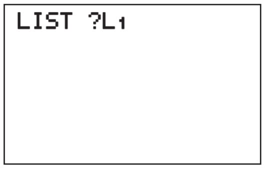
Finding a Confidence Interval for the Variance and Standard Deviation (Data)

1. Enter the data values into L_1.
2. Press **PRGM,** move the cursor to the program named SDINT, and press **ENTER** twice.
3. Press **1** for Data.
4. Type $\mathbf{L_1}$ for the list and press **ENTER.**
5. Type the confidence level and press **ENTER.**
6. Press **ENTER** to clear the screen.

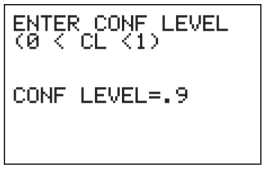
Example TI7–4

Find the 90% confidence interval for the variance and standard deviation for the data:

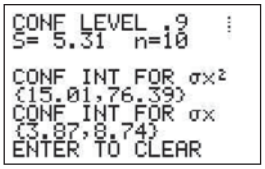
59 54 53 52 51 39 49 46 49 48



```
LIST ?L₁
```



```
ENTER CONF LEVEL
(0 < CL <1)

CONF LEVEL=.9
```



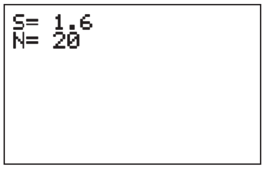
```
CONF LEVEL .9
S= 5.31  n=10

CONF INT FOR σx²
(15.01,76.39)
CONF INT FOR σx
(3.87,8.74)
ENTER TO CLEAR
```

Finding a Confidence Interval for the Variance and Standard Deviation (Statistics)

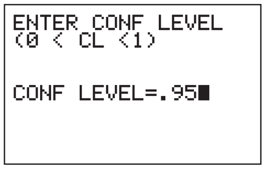
1. Press **PRGM,** move the cursor to the program named SDINT, and press **ENTER** twice.
2. Press **2** for Stats.
3. Type the sample standard deviation and press **ENTER.**
4. Type the sample size and press **ENTER.**
5. Type the confidence level and press **ENTER.**
6. Press **ENTER** to clear the screen.

Example TI7–5

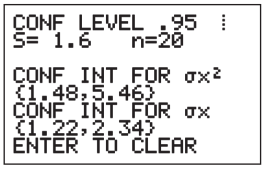
This refers to Example 7–14 in the text. Find the 95% confidence interval for the variance and standard deviation, given $n = 20$ and $s = 1.6$.



```
S= 1.6
N= 20
```



```
ENTER CONF LEVEL
(0 < CL <1)

CONF LEVEL=.95█
```



```
CONF LEVEL .95  ⋮
S= 1.6   n=20
CONF INT FOR σx²
{1.48,5.46}
CONF INT FOR σx
{1.22,2.34}
ENTER TO CLEAR
```

Summary

- An important aspect of inferential statistics is estimation. Estimations of parameters of populations are accomplished by selecting a random sample from that population and choosing and computing a statistic that is the best estimator of the parameter. A good estimator must be unbiased, consistent, and relatively efficient. The best estimate of μ is $\overline{X}$. (7–1)
- There are two types of estimates of a parameter: point estimates and interval estimates. A point estimate is a specific value. For example, if a researcher wishes to estimate the average length of a certain adult fish, a sample of the fish is selected and measured. The mean of this sample is computed, for example, 3.2 centimeters. From this sample mean, the researcher estimates the population mean to be 3.2 centimeters. The problem with point estimates is that the accuracy of the estimate cannot be determined. For this reason, statisticians prefer to use the interval estimate. By computing an interval about the sample value, statisticians can be 95 or 99% (or some other percentage) confident that their estimate contains the true parameter. The confidence level is determined by the researcher. The higher the confidence level, the wider the interval of the estimate must be. For example, a 95% confidence interval of the true mean length of a certain species of fish might be

$$3.17 < \mu < 3.23$$

whereas the 99% confidence interval might be

$$3.15 < \mu < 3.25 \quad (7\text{–}1)$$

- When the population standard deviation is known, the z value is used to compute the confidence interval. (7–1)
- Closely related to computing confidence intervals is the determination of the sample size to make an estimate of the mean. This information is needed to determine the minimum sample size necessary.
  1. The degree of confidence must be stated.
  2. The population standard deviation must be known or be able to be estimated.
  3. The margin of error must be stated. (7–1)
- If the population standard deviation is unknown, the t value is used. When the sample size is less than 30, the population must be normally distributed. (7–2)
- Confidence intervals and sample sizes can also be computed for proportions by using the normal distribution. (7–3)
- Finally, confidence intervals for variances and standard deviations can be computed by using the chi-square distribution. (7–4)

Important Terms

assumptions 370
chi-square distribution 398
confidence interval 371
confidence level 371
consistent estimator 371
degrees of freedom 383
estimation 370
estimator 371
interval estimate 371
margin of error 372
point estimate 370
proportion 390
relatively efficient estimator 371
robust 373
t distribution 383
unbiased estimator 371

Important Formulas

Formula for the confidence interval of the mean when σ is known (when $n \geq 30$, s can be used if σ is unknown):

$$\overline{X} - z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \overline{X} + z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

Formula for the sample size for means:

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E}\right)^2$$

where E is the margin of error.

Formula for the confidence interval of the mean when σ is unknown:

$$\overline{X} - t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) < \mu < \overline{X} + t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

Formula for the confidence interval for a proportion:

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

where $\hat{p} = X/n$ and $\hat{q} = 1 - \hat{p}$.

Formula for the sample size for proportions:

$$n = \hat{p}\hat{q}\left(\frac{z_{\alpha/2}}{E}\right)^2$$

Formula for the confidence interval for a variance:

$$\frac{(n-1)s^2}{\chi^2_{\text{right}}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{\text{left}}}$$

Formula for the confidence interval for a standard deviation:

$$\sqrt{\frac{(n-1)s^2}{\chi^2_{\text{right}}}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi^2_{\text{left}}}}$$

Review Exercises

Section 7–1

1. **Eye Blinks** A survey of 49 individuals found that the average number of times a person blinks his or her eyes is 25 times per minute. Based on this survey, find the 90% confidence of the population mean number of eye blinks per minute. The population standard deviation is 4.

2. **Vacation Days** A U.S. Travel Data Center survey reported that Americans stayed an average of 7.5 nights when they went on vacation. The sample size was 1500. Find a point estimate of the population mean. Find the 95% confidence interval of the true mean. Assume the population standard deviation was 0.8.

   Source: USA TODAY.

3. **Blood Pressure** A researcher wishes to estimate within 2 points the average systolic blood pressure of female college students. If she wishes to be 90% confident, how large a sample should she select if the population standard deviation of female systolic blood pressure is 4.8?

4. **Shopping Survey** A random sample of 49 shoppers showed that they spend an average of \$23.45 per visit at the Saturday Mornings Bookstore. The standard deviation of the population is \$2.80. Find a point estimate of the population mean. Find the 90% confidence interval of the true mean.

Section 7–2

5. **Lengths of Children's Animated Films** The lengths (in minutes) of a random selection of popular children's animated films are listed below. Estimate the true mean length of all children's animated films with 95% confidence. Assume the variable is normally distributed.

   93 83 76 92 77 81 78 100 78 76 75

6. **Dog Bites to Postal Workers** For a certain urban area, in a random sample of 5 months, an average of 28 mail carriers were bitten by dogs each month. The standard deviation of the sample was 3. Find the 90% confidence interval of the true mean number of mail carriers who are bitten by dogs each month. Assume the variable is normally distributed.

Section 7–3

7. **Number of Credit Cards** In a recent survey of 1000 adults ages 18 to 40, 34% said they had no credit cards. Find the 95% confidence interval of the population proportion.

   Source: CRC International.

8. **Vacation Sites** A U.S. Travel Data Center's survey of 1500 randomly selected adults found that 42% of respondents stated that they favor historical sites as vacations. Find the 95% confidence interval of the true proportion of all adults who favor visiting historical sites for vacations.

   Source: USA TODAY.

9. **Marriages in the United States** In 1960, 67.6% of the U.S. population (aged 14 and over) was married. A random sample of 600 Americans in 2014 indicated that 316 were married. Estimate the true proportion of

married Americans with 99% confidence. Assume the variable is normally distributed.

Source: World Almanac 2016.

10. **Adult Educational Activities** A local county has a very active adult education venue. A random sample of the population showed that 189 out of 400 persons 16 years old or older participated in some type of formal adult education activities, such as basic skills training, apprenticeships, personal interest courses, and part-time college or university degree programs. Estimate the true proportion of adults participating in some kind of formal education program with 98% confidence. Assume the variable is normally distributed.

11. **Cat Owners** A person recently read that 84% of cat owners are women. How large a sample should the person take if she wishes to be 90% confident that her proportion is within 3% of the true population proportion?

12. **Child Care Programs** A study found that 73% of randomly selected prekindergarten children ages 3 to 5 whose mothers had a bachelor's degree or higher were enrolled in center-based early childhood care and education programs. How large a sample is needed to estimate the true proportion within 3 percentage points with 95% confidence? How large a sample is needed if you had no prior knowledge of the proportion?

Section 7–4

13. **Baseball Diameters** The standard deviation of the diameter of 18 randomly selected baseballs was 0.29 cm. Find the 95% confidence interval of the true standard deviation of the diameters of the baseballs. Do you think the manufacturing process should be checked for inconsistency? Assume the variable is normally distributed.

14. **MPG for Lawn Mowers** A random sample of 22 lawn mowers was selected, and the motors were tested to see how many miles per gallon of gasoline each one obtained. The variance of the measurements was 2.6. Find the 95% confidence interval of the true variance. Assume the variable is normally distributed.

15. **Lifetimes of Snowmobiles** A random sample of 15 snowmobiles was selected, and the lifetime (in months) of the batteries was measured. The variance of the sample was 8.6. Find the 90% confidence interval of the true variance. Assume the variable is normally distributed.

16. **Length of Children's Animated Films** Use the data from Exercise 5 to estimate the population variance and standard deviation in length of children's animated films with 99% confidence. Assume the variable is normally distributed.

STATISTICS TODAY

Stress and the College Student —Revisited

The estimates given in the survey were point estimates; however, since the margin of error was 3 percentage points, a confidence interval can be constructed. For example, a confidence interval for the 85% figure would be 82% < p < 88%. We don't know whether this is a 90%, 95%, 99%, or some other confidence level because this was not specified in the report.

Using the formula in Section 7–3, a minimum sample size could be found. We can take $\hat{p}$ and $\hat{q}$ as 0.5 since no value for $\hat{p}$ is known, and $z_{\alpha/2} = 1.96$ for a 95% confidence interval.

$$n = \hat{p}\hat{q}\left(\frac{z_{\alpha/2}}{E}\right)^2$$

$$= (0.5)(0.5)\left(\frac{1.96}{0.03}\right)^2 = 1067.1$$

For a 95% confidence interval, then, the minimum sample size would be 1068 people (rounded up). Since the survey used 2240 students, it can be said that we are at least 95% confident that the true proportion of students who experience stress in college is somewhere between 82 and 88%.

Data Analysis

The Data Bank is found in Appendix B, or on the World Wide Web by following links from www.mhhe.com/math/stat/bluman/.

1. From the Data Bank choose a variable, find the mean, and construct the 95 and 99% confidence intervals of the population mean. Use a sample of at least 30 subjects. Find the mean of the population, and determine whether it falls within the confidence interval.
2. Repeat Exercise 1, using a different variable and a sample of 15.
3. Repeat Exercise 1, using a proportion. For example, construct a confidence interval for the proportion of individuals who did not complete high school.
4. From Data Set III in Appendix B, select a sample of 30 values and construct the 95 and 99% confidence intervals of the mean length in miles of major North American rivers. Find the mean of all the values, and determine if the confidence intervals contain the mean.
5. From Data Set VI in Appendix B, select a sample of 20 values and find the 90% confidence interval of the mean of the number of acres. Find the mean of all the values, and determine if the confidence interval contains the mean.
6. Select a random sample of 20 of the record high temperatures in the United States, found in Data Set I in Appendix B. Find the proportion of temperatures below 110°. Construct a 95% confidence interval for this proportion. Then find the true proportion of temperatures below 110°, using all the data. Is the true proportion contained in the confidence interval? Explain.

Chapter Quiz

Determine whether each statement is true or false. If the statement is false, explain why.

1. Interval estimates are preferred over point estimates since a confidence level can be specified.
2. For a specific confidence interval, the larger the sample size, the smaller the margin of error will be.
3. An estimator is consistent if as the sample size decreases, the value of the estimator approaches the value of the parameter estimated.
4. To determine the sample size needed to estimate a parameter, you must know the margin of error.

Select the best answer.

5. When a 99% confidence interval is calculated instead of a 95% confidence interval with n being the same, the margin of error will be
   a. Smaller *c.* The same
   b. Larger *d.* It cannot be determined.
6. The best point estimate of the population mean is
   a. The sample mean
   b. The sample median
   c. The sample mode
   d. The sample midrange
7. When the population standard deviation is unknown and the sample size is less than 30, what table value should be used in computing a confidence interval for a mean?
   a. z
   b. t
   c. Chi-square
   d. None of the above

Complete the following statements with the best answer.

8. A good estimator should be _______, _______, and _______.
9. The maximum likely difference between the point estimate of a parameter and the actual value of the parameter is called the _______.
10. The statement "The average height of an adult male is 5 feet 10 inches" is an example of a(n) _______ estimate.
11. The three confidence levels used most often are the _______%, _______%, and _______%.
12. **Cost of Texts** An irate student complained that the cost of textbooks was too high. He randomly surveyed 36 other students and found that the mean amount of money spent for texts was \$121.60. If the standard deviation of the population was \$6.36, find the best point estimate and the 90% confidence interval of the true mean.
13. **Doctor Visit Costs** An irate patient complained that the cost of a doctor's visit was too high. She randomly surveyed 20 other patients and found that the mean amount of money they spent on each doctor's visit was \$44.80. The standard deviation of the sample was \$3.53. Find a point estimate of the population mean. Find the 95% confidence interval of the population mean. Assume the variable is normally distributed.

14. **Weights of Minivans** The average weight of 40 randomly selected minivans was 4150 pounds. The population standard deviation was 480 pounds. Find a point estimate of the population mean. Find the 99% confidence interval of the true mean weight of the minivans.

15. **Ages of Insurance Representatives** In a study of 10 randomly selected insurance sales representatives from a certain large city, the average age of the group was 48.6 years and the standard deviation was 4.1 years. Assume the variable is normally distributed. Find the 95% confidence interval of the population mean age of all insurance sales representatives in that city.

16. **Patients Treated in Hospital Emergency Rooms** In a hospital, a random sample of 8 weeks was selected, and it was found that an average of 438 patients were treated in the emergency room each week. The standard deviation was 16. Find the 99% confidence interval of the true mean. Assume the variable is normally distributed.

17. **Burglaries** For a certain urban area, it was found that in a random sample of 4 months, an average of 31 burglaries occurred each month. The standard deviation was 4. Assume the variable is normally distributed. Find the 90% confidence interval of the true mean number of burglaries each month.

18. **Hours Spent Studying** A university dean wishes to estimate the average number of hours that freshmen study each week. The standard deviation from a previous study is 2.6 hours. How large a sample must be selected if he wants to be 99% confident of finding whether the true mean differs from the sample mean by 0.5 hour?

19. **Money Spent on Road Repairs** A researcher wishes to estimate within $300 the true average amount of money a county spends on road repairs each year. If she wants to be 90% confident, how large a sample is necessary? The standard deviation is known to be $900.

20. **Political Survey** A political analyst found that 43% of 300 randomly selected Republican voters feel that the federal government has too much power. Find the 95% confidence interval of the population proportion of Republican voters who feel this way.

21. **Emergency Room Accidents** In a study of 150 randomly selected accidents that required treatment in an emergency room, 36% involved children under 6 years of age. Find the 90% confidence interval of the true proportion of accidents that involve children under the age of 6.

22. **Television Set Ownership** A survey of 90 randomly selected families showed that 40 owned at least one television set. Find the 95% confidence interval of the true proportion of families who own at least one television set.

23. **Skipping Lunch** A nutritionist wishes to determine, within 3%, the true proportion of adults who do not eat any lunch. If he wishes to be 95% confident that his estimate contains the population proportion, how large a sample will be necessary? A previous study found that 15% of the 125 people surveyed said they did not eat lunch.

24. **Novel Pages** A random sample of 25 novels has a standard deviation of 9 pages. Find the 95% confidence interval of the population standard deviation. Assume the variable is normally distributed.

25. **Truck Safety Check** Find the 90% confidence interval for the variance and standard deviation for the time it takes a state police inspector to check a truck for safety if a random sample of 27 trucks has a standard deviation of 6.8 minutes. Assume the variable is normally distributed.

26. **Automobile Pollution** A random sample of 20 automobiles has a pollution by-product release standard deviation of 2.3 ounces when 1 gallon of gasoline is used. Find the 90% confidence interval of the population standard deviation. Assume the variable is normally distributed.

Critical Thinking Challenges

A confidence interval for a median can be found by using these formulas

$$U = \frac{n+1}{2} + \frac{z_{\alpha/2}\sqrt{n}}{2} \qquad \text{(round up)}$$

$$L = n - U + 1$$

to define positions in the set of ordered data values.

Suppose a data set has 30 values, and you want to find the 95% confidence interval for the median. Substituting in the formulas, you get

$$U = \frac{30+1}{2} + \frac{1.96\sqrt{30}}{2} = 21 \qquad \text{(rounded up)}$$

$$L = 30 - 21 + 1 = 10$$

when $n = 30$ and $z_{\alpha/2} = 1.96$.

Arrange the data in order from smallest to largest, and then select the 10th and 21st values of the data array; hence, $X_{10} < \text{median} < X_{21}$.

Using the previous information, find the 90% confidence interval for the median for the given data.

84	49	3	133	85	4340	461	60	28	97
14	252	18	16	24	346	254	29	254	6
31	104	72	29	391	19	125	10	6	17
72	31	23	225	72	5	61	366	77	8
26	8	55	138	158	846	123	47	21	82

Data Projects

1. **Business and Finance** Use 30 stocks classified as the Dow Jones industrials as the sample. Note the amount each stock has gained or lost in the last quarter. Compute the mean and standard deviation for the data set. Compute the 95% confidence interval for the mean and the 95% confidence interval for the standard deviation. Compute the percentage of stocks that had a gain in the last quarter. Find a 95% confidence interval for the percentage of stocks with a gain.
2. **Sports and Leisure** Use the top home run hitter from each major league baseball team as the data set. Find the mean and the standard deviation for the number of home runs hit by the top hitter on each team. Find a 95% confidence interval for the mean number of home runs hit.
3. **Technology** Use the data collected in data project 3 of Chapter 2 regarding song lengths. Select a specific genre, and compute the percentage of songs in the sample that are of that genre. Create a 95% confidence interval for the true percentage. Use the entire music library, and find the population percentage of the library with that genre. Does the population percentage fall within the confidence interval?
4. **Health and Wellness** Use your class as the sample. Have each student take her or his temperature on a day when he or she is healthy. Compute the mean and standard deviation for the sample. Create a 95% confidence interval for the mean temperature. Does the confidence interval obtained support the long-held belief that the average body temperature is 98.6°F?
5. **Politics and Economics** Select five political polls and note the margin of error, sample size, and percentage favoring the candidate for each. For each poll, determine the level of confidence that must have been used to obtain the margin of error given, knowing the percentage favoring the candidate and number of participants. Is there a pattern that emerges?
6. **Your Class** Have each student compute his or her body mass index (BMI) (703 times weight in pounds, divided by the square of the height in inches). Find the mean and standard deviation for the data set. Compute a 95% confidence interval for the mean BMI of a student. A BMI score over 30 is considered obese. Does the confidence interval indicate that the mean for BMI could be in the obese range?

Answers to Applying the Concepts

Section 7–1 How Many Tissues Should Be in a Box?

1. Answers will vary. One possible answer is to find out the average number of tissues that a group of randomly selected individuals use in a 2-week period.
2. People usually need tissues when they have colds or when their allergies are acting up.
3. If we want to concentrate on the number of tissues used when people have colds, we select a random sample of people with colds and have them keep a record of how many tissues they use during their colds.
4. Answers may vary. I will use a 95% confidence interval:

$$\bar{x} \pm 1.96\frac{\sigma}{\sqrt{n}} = 57 \pm 1.96\frac{15}{\sqrt{85}} = 57 \pm 3$$

   I am 95% confident that the interval 54–60 contains the true mean number of tissues used by people when they have colds. It seems reasonable to put 60 tissues in the new automobile glove compartment boxes.
5. Answers will vary. Since I am 95% confident that the interval contains the true average, any number of tissues between 54 and 60 would be reasonable. Sixty seemed to be the most reasonable answer, since it is close to 2 standard deviations above the sample mean.

Section 7–2 Sport Drink Decision

1. Answers will vary. One possible answer is that this is a small sample since I am only looking at seven popular sport drinks.
2. The mean cost per container is $1.25, with standard deviation of $0.39. The 90% confidence interval is

$$\bar{X} \pm t_{\alpha/2}\frac{s}{\sqrt{n}} = 1.25 \pm 1.943\frac{0.39}{\sqrt{7}} = 1.25 \pm 0.29$$

   or $0.96 < \mu < 1.54$

   The 10-K, All Sport, Exceed, and Hydra Fuel all fall outside of the confidence interval.
3. Answers will vary. $1.89 is far enough outside the interval that it may affect its results.
4. There are $7 - 1 = 6$ degrees of freedom.
5. Cost per serving would impact my decision on purchasing a sport drink, since this would allow me to compare the costs on an equal scale.
6. Answers will vary.

Section 7–3 Contracting Influenza

1. (95% CI) means that these are the 95% confidence intervals constructed from the data.
2. The margin of error for men reporting influenza is $(50.5 - 47.1)/2 = 1.7\%$.

3. The total sample size was 19,774.
4. The larger the sample size, the smaller the margin of error and the narrower the confidence interval (all other things being held constant).
5. A 90% confidence interval would be narrower (smaller) than a 95% confidence interval, since we need to include fewer values in the interval.
6. The 51.5% is the middle of the confidence interval, since it is the point estimate for the confidence interval.

Section 7–4 Ages of Presidents at the Time of Their Deaths

1. The data represent a population, since we have the age at death for all deceased Presidents (at the time of the writing of this book).
2. Answers will vary. One possible sample is 56, 67, 53, 46, 63, 77, 63, 57, 71, 57, 80, 65, which results in a standard deviation of 9.9 years and a variance of 98.1.
3. Answers will vary. The 95% confidence interval for the standard deviation is $\sqrt{\frac{(n-1)s^2}{\chi^2_{\text{right}}}}$ to $\sqrt{\frac{(n-1)s^2}{\chi^2_{\text{left}}}}$. In this case we have $\sqrt{\frac{(12-1)9.9^2}{21.920}} = \sqrt{49.1839} = 7.0$ to $\sqrt{\frac{(12-1)9.9^2}{3.816}} = \sqrt{282.524} = 16.8$, or 7.0 to 16.8 years.
4. The standard deviation for all the data values is 11.9 years.
5. Answers will vary. Yes, the confidence interval does contain the population standard deviation.
6. Answers will vary.
7. We need to assume that the distribution of ages at death is normal.

8

Hypothesis Testing

© Photosindia.Com, Llc/Glow Images RF

STATISTICS TODAY

How Much Better Is Better?

Suppose a school superintendent reads an article which states that the overall mean score for the SAT is 910. Furthermore, suppose that, for a sample of students, the average of the SAT scores in the superintendent's school district is 960. Can the superintendent conclude that the students in his school district scored higher on average? At first glance, you might be inclined to say yes, since 960 is higher than 910. But recall that the means of samples vary about the population mean when samples are selected from a specific population. So the question arises, is there a real difference in the means, or is the difference simply due to chance (i.e., sampling error)? In this chapter, you will learn how to answer that question by using statistics that explain hypothesis testing. See Statistics Today—Revisited at the end of the chapter for the answer. In this chapter, you will learn how to answer many questions of this type by using statistics that are explained in the theory of hypothesis testing.

OUTLINE

Introduction
8–1 Steps in Hypothesis Testing—Traditional Method
8–2 z Test for a Mean
8–3 t Test for a Mean
8–4 z Test for a Proportion
8–5 χ^2 Test for a Variance or Standard Deviation
8–6 Additional Topics Regarding Hypothesis Testing
Summary

OBJECTIVES

After completing this chapter, you should be able to:

① Understand the definitions used in hypothesis testing.

② State the null and alternative hypotheses.

③ Find critical values for the z test.

④ State the five steps used in hypothesis testing.

⑤ Test means when σ is known, using the z test.

⑥ Test means when σ is unknown, using the t test.

⑦ Test proportions, using the z test.

⑧ Test variances or standard deviations, using the chi-square test.

⑨ Test hypotheses, using confidence intervals.

⑩ Explain the relationship between type I and type II errors and the power of a test.

Introduction

Researchers are interested in answering many types of questions. For example, a scientist might want to know whether there is evidence of global warming. A physician might want to know whether a new medication will lower a person's blood pressure. An educator might wish to see whether a new teaching technique is better than a traditional one. A retail merchant might want to know whether the public prefers a certain color in a new line of fashion. Automobile manufacturers are interested in determining whether a new type of seat belt will reduce the severity of injuries caused by accidents. These types of questions can be addressed through statistical **hypothesis testing,** which is a decision-making process for evaluating claims about a population. In hypothesis testing, the researcher must define the population under study, state the particular hypotheses that will be investigated, give the significance level, select a sample from the population, collect the data, perform the calculations required for the statistical test, and reach a conclusion.

Hypotheses concerning parameters such as means and proportions can be investigated. There are two specific statistical tests used for hypotheses concerning means: the *z test* and the *t test.* This chapter will explain in detail the hypothesis-testing procedure along with the z test and the t test. In addition, a hypothesis-testing procedure for testing a single variance or standard deviation using the chi-square distribution is explained in Section 8–5.

The three methods used to test hypotheses are

1. The traditional method
2. The P-value method
3. The confidence interval method

The *traditional method* will be explained first. It has been used since the hypothesis-testing method was formulated. A newer method, called the *P-value method,* has become popular with the advent of modern computers and high-powered statistical calculators. It will be explained at the end of Section 8–2. The third method, the *confidence interval method,* is explained in Section 8–6 and illustrates the relationship between hypothesis testing and confidence intervals.

8–1 Steps in Hypothesis Testing—Traditional Method

Every hypothesis-testing situation begins with the statement of a hypothesis.

A **statistical hypothesis** is a conjecture about a population parameter. This conjecture may or may not be true.

OBJECTIVE ❶

Understand the definitions used in hypothesis testing.

There are two types of statistical hypotheses for each situation: the null hypothesis and the alternative hypothesis.

The **null hypothesis,** symbolized by H_0, is a statistical hypothesis that states that there is no difference between a parameter and a specific value, or that there is no difference between two parameters.

The **alternative hypothesis,** symbolized by H_1, is a statistical hypothesis that states the existence of a difference between a parameter and a specific value, or states that there is a difference between two parameters.

(*Note:* Although the definitions of null and alternative hypotheses given here use the word *parameter,* these definitions can be extended to include other terms such as *distributions* and *randomness*. This is explained in later chapters.)

As an illustration of how hypotheses should be stated, three different statistical studies will be used as examples.

OBJECTIVE 2

State the null and alternative hypotheses.

Situation A A medical researcher is interested in finding out whether a new medication will have any undesirable side effects. The researcher is particularly concerned with the pulse rate of the patients who take the medication. Will the pulse rate increase, decrease, or remain unchanged after a patient takes the medication?

Since the researcher knows that the mean pulse rate for the population under study is 82 beats per minute, the hypotheses for this situation are

$$H_0\colon \mu = 82 \qquad \text{and} \qquad H_1\colon \mu \neq 82$$

The null hypothesis specifies that the mean will remain unchanged, and the alternative hypothesis states that it will be different. This test is called a *two-tailed test* (a term that will be formally defined later in this section), since the possible side effects of the medicine could be to raise or lower the pulse rate.

Situation B A chemist invents an additive to increase the life of an automobile battery. If the mean lifetime of the automobile battery without the additive is 36 months, then her hypotheses are

$$H_0\colon \mu = 36 \qquad \text{and} \qquad H_1\colon \mu > 36$$

In this situation, the chemist is interested only in increasing the lifetime of the batteries, so her alternative hypothesis is that the mean is greater than 36 months. The null hypothesis is that the mean is equal to 36 months. This test is called *right-tailed,* since the interest is in an increase only.

Unusual Stat

Sixty-three percent of people would rather hear bad news before hearing the good news.

Situation C A contractor wishes to lower heating bills by using a special type of insulation in houses. If the average of the monthly heating bills is \$78, her hypotheses about heating costs with the use of insulation are

$$H_0\colon \mu = \$78 \qquad \text{and} \qquad H_1\colon \mu < \$78$$

This test is a *left-tailed test,* since the contractor is interested only in lowering heating costs.

To state hypotheses correctly, researchers must translate the *conjecture* or *claim* from words into mathematical symbols. The basic symbols used are as follows:

Equal to	$=$	Greater than	$>$
Not equal to	$\neq$	Less than	$<$

The null and alternative hypotheses are stated together, and the null hypothesis contains the equals sign, as shown (where k represents a specified number).

Also notice that in a right-tailed test, the inequality sign points to the right; and in a left-tailed test, the inequality sign points to the left. Right- and left-tailed tests are also called one-tailed tests.

Two-tailed test	Right-tailed test	Left-tailed test
$H_0\colon \mu = k$	$H_0\colon \mu = k$	$H_0\colon \mu = k$
$H_1\colon \mu \neq k$	$H_1\colon \mu > k$	$H_1\colon \mu < k$

The formal definitions of the different types of tests are given later in this section.

In this book, the null hypothesis is always stated using the equals sign. This is done because in most professional journals, and when we test the null hypothesis, the assumption is that the mean, proportion, or standard deviation is equal to a given specific value. Also, when a researcher conducts a study, he or she is generally looking for evidence to support a claim. Therefore, the claim should be stated as the alternative hypothesis, i.e., using $<$ or $>$ or $\neq$. Because of this, the alternative hypothesis is sometimes called the **research hypothesis.**

TABLE 8–1 Hypothesis-Testing Common Phrases

$>$	$<$
Is greater than	Is less than
Is above	Is below
Is higher than	Is lower than
Is longer than	Is shorter than
Is bigger than	Is smaller than
Is increased	Is decreased or reduced from
$=$	$\neq$
Is equal to	Is not equal to
Is no different from	Is different from
Has not changed from	Has changed from
Is the same as	Is not the same as

A claim, however, can be stated as either the null hypothesis or the alternative hypothesis; but the statistical evidence can only *support* the claim if it is the alternative hypothesis. Statistical evidence can be used to *reject* the claim if the claim is the null hypothesis. These facts are important when you are stating the conclusion of a statistical study.

Table 8–1 shows some common phrases that are used in hypotheses and conjectures, and the corresponding symbols. This table should be helpful in translating verbal conjectures into mathematical symbols.

EXAMPLE 8–1

State the null and alternative hypotheses for each conjuncture.

a. A researcher studies gambling in young people. She thinks those who gamble spend more than $30 per day.

b. A researcher wishes to see if police officers whose spouses work in law enforcement have a lower score on a work stress questionnaire than the average score of 120.

c. A teacher feels that if an online textbook is used for a course instead of a hardback book, it may change the students' scores on a final exam. In the past, the average final exam score for the students was 83.

SOLUTION

a. H_0: $\mu = \$30$ and H_1: $\mu > \$30$

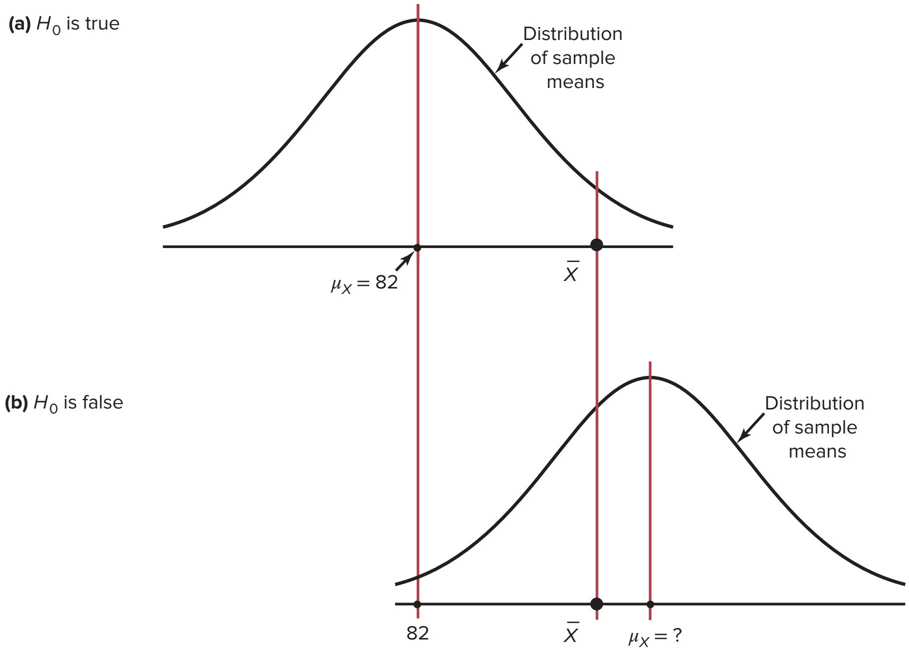
b. H_0: $\mu = 120$ and H_1: $\mu < 120$

c. H_0: $\mu = 83$ and H_1: $\mu \neq 83$

After stating the hypothesis, the researcher designs the study. The researcher selects the correct *statistical test,* chooses an appropriate *level of significance,* and formulates a plan for conducting the study. In situation A, for instance, the researcher will select a sample of patients who will be given the drug. After allowing a suitable time for the drug to be absorbed, the researcher will measure each person's pulse rate.

Recall that when samples of a specific size are selected from a population, the means of these samples will vary about the population mean, and the distribution of the sample means will be approximately normal when the sample size is 30 or more. (See Section 6–3.) So even if the null hypothesis is true, the mean of the pulse rates of the sample of patients will not, in most cases, be exactly equal to the population mean of 82 beats per minute. There are

FIGURE 8–1
Situations in Hypothesis Testing



two possibilities. Either the null hypothesis is true, and the difference between the sample mean and the population mean is due to chance; *or* the null hypothesis is false, and the sample came from a population whose mean is not 82 beats per minute but is some other value that is not known. These situations are shown in Figure 8–1.

The farther away the sample mean is from the population mean, the more evidence there would be for rejecting the null hypothesis. The probability that the sample came from a population whose mean is 82 decreases as the distance or absolute value of the difference between the means increases.

If the mean pulse rate of the sample were, say, 83, the researcher would probably conclude that this difference was due to chance and would not reject the null hypothesis. But if the sample mean were, say, 90, then in all likelihood the researcher would conclude that the medication increased the pulse rate of the users and would reject the null hypothesis. The question is, Where does the researcher draw the line? This decision is not made on feelings or intuition; it is made statistically. That is, the difference must be significant and in all likelihood not due to chance. Here is where the concepts of statistical test and level of significance are used.

A **statistical test** uses the data obtained from a sample to make a decision about whether the null hypothesis should be rejected.

The numerical value obtained from a statistical test is called the **test value or test statistic.**

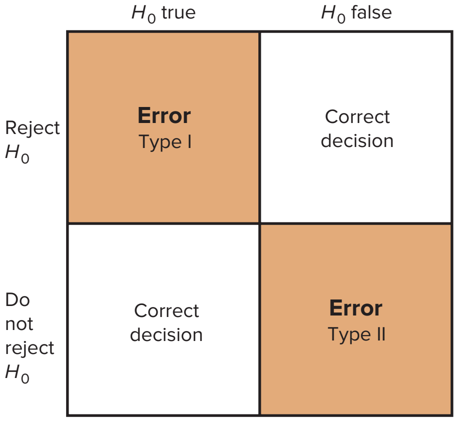
In this type of statistical test, the mean is computed for the data obtained from the sample and is compared with the population mean. Then a decision is made to reject or not reject the null hypothesis on the basis of the value obtained from the statistical test. If the difference is significant, the null hypothesis is rejected. If it is not, then the null hypothesis is not rejected.

As previously stated, sample data are used to determine if a null hypothesis should be rejected. Because this decision is based on sample data, there is a possibility that an incorrect decision can be made.

In the hypothesis-testing situation, there are four possible outcomes. In reality, the null hypothesis may or may not be true, and a decision is made to reject or not reject it on

FIGURE 8–2

Possible Outcomes of a Hypothesis Test



the basis of the data obtained from a sample. The four possible outcomes are shown in Figure 8–2. Notice that there are two possibilities for a correct decision and two possibilities for an incorrect decision.

The four possibilities are as follows:

1. We reject the null hypothesis when it is true. This would be an incorrect decision and would result in a **type I error.**
2. We reject the null hypothesis when it is false. This would be a correct decision.
3. We do not reject the null hypothesis when it is true. This would be a correct decision.
4. We do not reject the null hypothesis when it is false. This would be an incorrect decision and would result in a **type II error.**

A **type I error** occurs if you reject the null hypothesis when it is true.

A **type II error** occurs if you do not reject the null hypothesis when it is false.

If a null hypothesis is true and it is rejected, then a *type I error* is made. In situation A, for instance, the medication might not significantly change the pulse rate of all the users in the population; but it might change the rate, by chance, of the subjects in the sample. In this case, the researcher will reject the null hypothesis when it is really true, thus committing a type I error.

On the other hand, the medication might not change the pulse rate of the subjects in the sample; but when it is given to the general population, it might cause a significant increase or decrease in the pulse rate of users. The researcher, on the basis of the data obtained from the sample, will not reject the null hypothesis, thus committing a *type II error.*

In situation B, the additive might not significantly increase the lifetimes of automobile batteries in the population, but it might increase the lifetimes of the batteries in the sample. In this case, the null hypothesis would be rejected when it was really true. This would be a type I error. On the other hand, the additive might not work on the batteries selected for the sample, but if it were to be used in the general population of batteries, it might significantly increase their lifetimes. The researcher, on the basis of information obtained from the sample, would not reject the null hypothesis, thus committing a type II error.

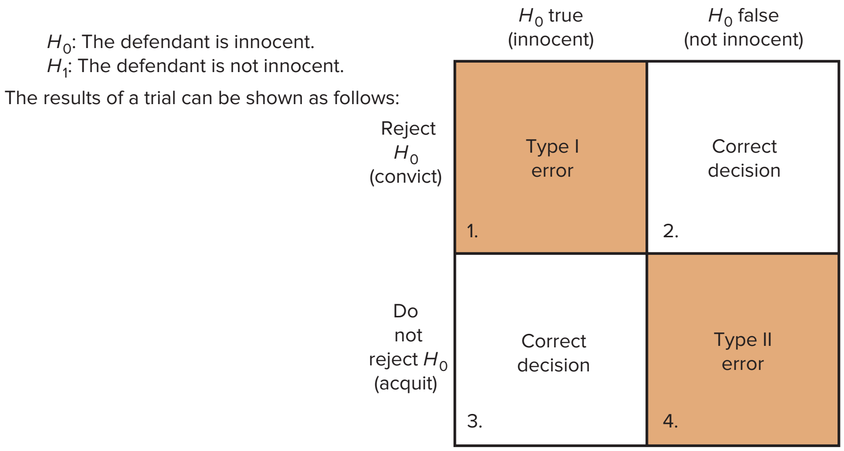
The hypothesis-testing situation can be likened to a jury trial. In a jury trial, there are four possible outcomes. The defendant is either guilty or innocent, and he or she will be convicted or acquitted. See Figure 8–3.

Now the hypotheses are

H_0: The defendant is innocent.

H_1: The defendant is not innocent (i.e., guilty).

FIGURE 8–3
Hypothesis Testing and a Jury Trial



H_0: The defendant is innocent.
H_1: The defendant is not innocent.

The results of a trial can be shown as follows:

	H_0 true (innocent)	H_0 false (not innocent)
Reject H_0 (convict)	1. Type I error	2. Correct decision
Do not reject H_0 (acquit)	3. Correct decision	4. Type II error

Next, the evidence is presented in court by the prosecutor, and based on this evidence, the jury decides the verdict, guilty or not guilty.

If the defendant is convicted but he or she did not commit the crime, then a type I error has been committed. See block 1 of Figure 8–3. In other words, a type I error means an innocent person is put in jail. On the other hand, if the defendant is convicted and he or she has committed the crime, then a correct decision has been made. See block 2.

If the defendant is acquitted and he or she did not commit the crime, a correct decision has been made by the jury. See block 3. However, if the defendant is acquitted and he or she did commit the crime, then a type II error has been made. See block 4. In other words, a type II error is letting a guilty person go free.

The decision of the jury does not prove that the defendant did or did not commit the crime. The decision is based on the evidence presented. If the evidence is strong enough, the defendant will be convicted in most cases. If the evidence is weak, the defendant will be acquitted in most cases. Nothing is proved absolutely. Likewise, the decision to reject or not reject the null hypothesis does not prove anything. *The only way to prove anything statistically is to use the entire population,* which, in most cases, is not possible. The decision, then, is made on the basis of probabilities. That is, when there is a large difference between the mean obtained from the sample and the hypothesized mean, the null hypothesis is probably not true. The question is, How large a difference is necessary to reject the null hypothesis? Here is where the level of significance is used.

Unusual Stats

Of workers in the United States, 64% drive to work alone and 6% of workers walk to work.

The **level of significance** is the maximum probability of committing a type I error. This probability is symbolized by α (Greek letter **alpha**). That is, $P(\text{type I error}) = \alpha$.

The probability of a type II error is symbolized by β, the Greek letter **beta.** That is, $P(\text{type II error}) = \beta$. In most hypothesis-testing situations, β cannot be easily computed; however, α and β are related in that decreasing one increases the other.

Statisticians generally agree on using three arbitrary significance levels: the 0.10, 0.05, and 0.01 levels. That is, if the null hypothesis is rejected, the probability of a type I error will be 10, 5, or 1%, depending on which level of significance is used. Here is another way of putting it: When $\alpha = 0.10$, there is a 10% chance of rejecting a true null hypothesis; when $\alpha = 0.05$, there is a 5% chance of rejecting a true null hypothesis; and when $\alpha = 0.01$, there is a 1% chance of rejecting a true null hypothesis.

In a hypothesis-testing situation, the researcher decides what level of significance to use. It does not have to be the 0.10, 0.05, or 0.01 level. It can be any level, depending on the seriousness of the type I error. After a significance level is chosen, a *critical value* is selected from a table for the appropriate test. If a z test is used, for example, the z table

(Table E in Appendix A) is consulted to find the critical value. The critical value determines the critical and noncritical regions.

The **critical** or **rejection region** is the range of test values that indicates that there is a significant difference and that the null hypothesis should be rejected.

The **noncritical** or **nonrejection region** is the range of test values that indicates that the difference was probably due to chance and that the null hypothesis should not be rejected.

The **critical value** separates the critical region from the noncritical region. The symbol for critical value is C.V.

The critical value can be on the right side of the mean or on the left side of the mean for a one-tailed test. Its location depends on the inequality sign of the alternative hypothesis. For example, in situation B, where the chemist is interested in increasing the average lifetime of automobile batteries, the alternative hypothesis is H_1: $\mu > 36$. Since the inequality sign is >, the null hypothesis will be rejected only when the sample mean is significantly greater than 36. Hence, the critical value must be on the right side of the mean. Therefore, this test is called a right-tailed test.

A **one-tailed test** indicates that the null hypothesis should be rejected when the test value is in the critical region on one side of the mean. A one-tailed test is either a **right-tailed test** or a **left-tailed test,** depending on the direction of the inequality of the alternative hypothesis.

To obtain the critical value, the researcher must choose an alpha level. In situation B, suppose the researcher chose $\alpha = 0.01$. Then the researcher must find a z value such that 1% of the area falls to the right of the z value and 99% falls to the left of the z value, as shown in Figure 8–4(a).

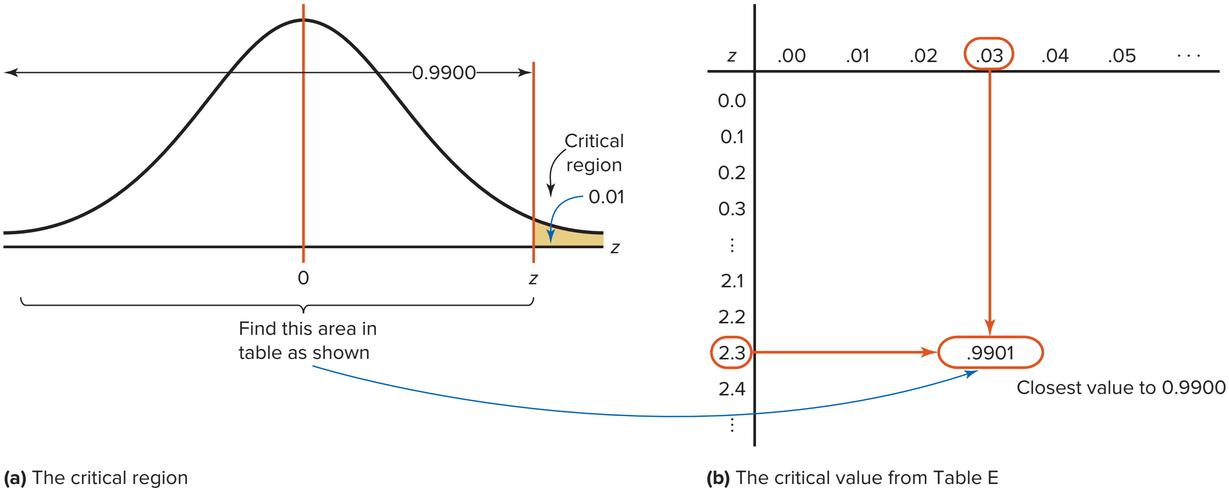
OBJECTIVE 3

Find critical values for the *z* test.

Next, the researcher must find the area value in Table E closest to 0.9900. The critical z value is 2.33, since that value gives the area closest to 0.9900 (that is, 0.9901), as shown in Figure 8–4(b).

The critical and noncritical regions and the critical value are shown in Figure 8–5.

FIGURE 8–4 Finding the Critical Value for $\alpha = 0.01$ (Right-Tailed Test)



(a) The critical region

(b) The critical value from Table E

FIGURE 8–5
Critical and Noncritical Regions for $\alpha = 0.01$ (Right-Tailed Test)

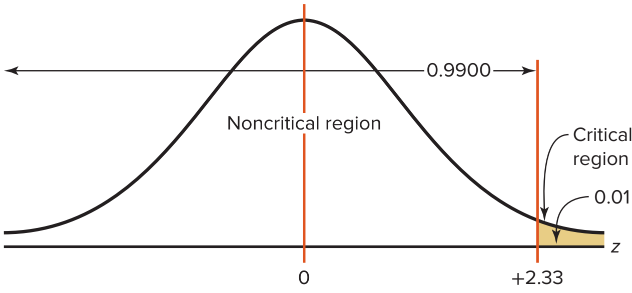
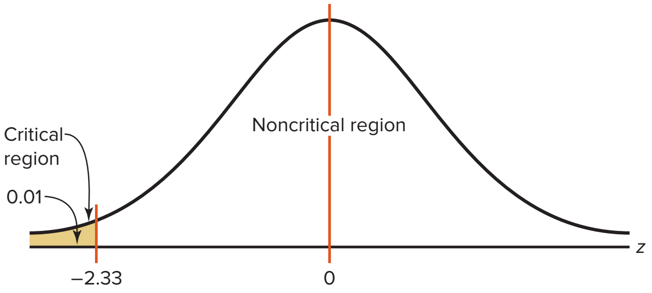


FIGURE 8–6
Critical and Noncritical Regions for $\alpha = 0.01$ (Left-Tailed Test)



Now, move on to situation C, where the contractor is interested in lowering the heating bills. The alternative hypothesis is H_1: $\mu <$ \$78. Hence, the critical value falls to the left of the mean. This test is thus a left-tailed test. At $\alpha = 0.01$, the critical value is −2.33, since 0.0099 is the closest value to 0.01. This is shown in Figure 8–6.

When a researcher conducts a two-tailed test, as in situation A, the null hypothesis can be rejected when there is a significant difference in either direction, above or below the mean.

In a **two-tailed test,** the null hypothesis should be rejected when the test value is in either of the two critical regions.

For a two-tailed test, then, the critical region must be split into two equal parts. If $\alpha = 0.01$, then one-half of the area, or 0.005, must be to the right of the mean and one-half must be to the left of the mean, as shown in Figure 8–7.

In this case, the z value on the left side is found by looking up the z value corresponding to an area of 0.0050. The z value falls about halfway between −2.57 and −2.58 corresponding to the areas 0.0051 and 0.0049. The average of −2.57 and −2.58 is $[(-2.57) + (-2.58)] \div 2 = -2.575$ so if the z value is needed to three decimal places, −2.575 is used; however, if the z value is rounded to two decimal places, −2.58 is used.

On the right side, it is necessary to find the z value corresponding to 0.99 + 0.005, or 0.9950. Again, the value falls between 0.9949 and 0.9951, so +2.575 or 2.58 can be used. See Figure 8–7.

FIGURE 8–7
Finding the Critical Values for $\alpha = 0.01$ (Two-Tailed Test)

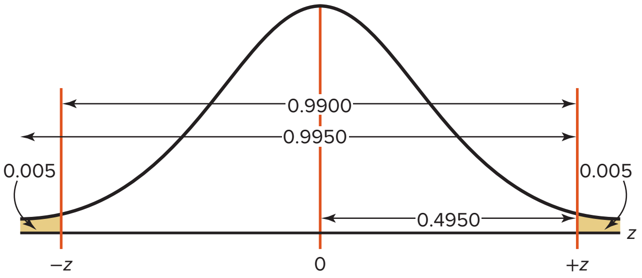


FIGURE 8–8
Critical and Noncritical Regions for $\alpha = 0.01$ (Two-Tailed Test)

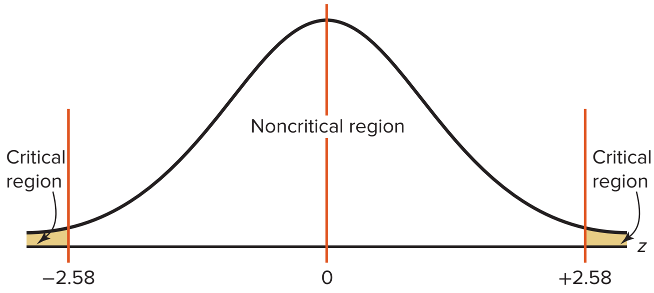
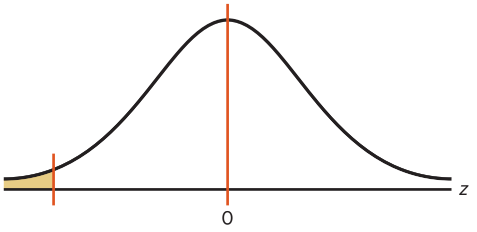


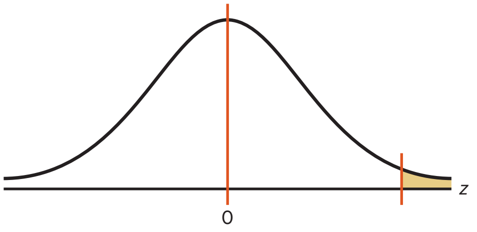
FIGURE 8–9
Summary of Hypothesis Testing and Critical Values

$$\begin{array}{l} H_0\!: \mu = k \\ H_1\!: \mu < k \end{array} \begin{cases} \alpha = 0.10,\ \text{C.V.} = -1.28 \\ \alpha = 0.05,\ \text{C.V.} = -1.65 \\ \alpha = 0.01,\ \text{C.V.} = -2.33 \end{cases}$$



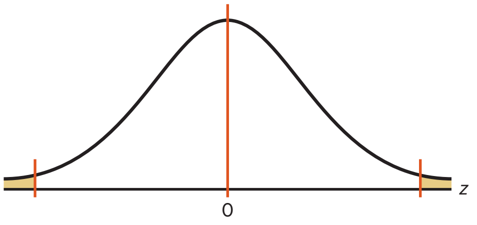
(a) Left-tailed

$$\begin{array}{l} H_0\!: \mu = k \\ H_1\!: \mu > k \end{array} \begin{cases} \alpha = 0.10,\ \text{C.V.} = +1.28 \\ \alpha = 0.05,\ \text{C.V.} = +1.65 \\ \alpha = 0.01,\ \text{C.V.} = +2.33 \end{cases}$$



(b) Right-tailed

$$\begin{array}{l} H_0\!: \mu = k \\ H_1\!: \mu \neq k \end{array} \begin{cases} \alpha = 0.10,\ \text{C.V.} = \pm 1.65 \\ \alpha = 0.05,\ \text{C.V.} = \pm 1.96 \\ \alpha = 0.01,\ \text{C.V.} = \pm 2.58 \end{cases}$$



(c) Two-tailed

The critical values are +2.58 and −2.58, as shown in Figure 8–8.

Similar procedures are used to find other values of α.

Figure 8–9 with rejection regions shaded shows the critical values (C.V.) for the three situations discussed in this section for $\alpha = 0.10$, $\alpha = 0.05$, and $\alpha = 0.01$. The procedure for finding critical values is outlined next (where k is a specified number).

Procedure Table

Finding the Critical Values for Specific α Values, Using Table E

Step 1 Draw the figure and indicate the appropriate area.

a. If the test is left-tailed, the critical region, with an area equal to α, will be on the left side of the mean.

b. If the test is right-tailed, the critical region, with an area equal to α, will be on the right side of the mean.

c. If the test is two-tailed, α must be divided by 2; one-half of the area will be to the right of the mean, and one-half will be to the left of the mean.

Step 2 *a.* For a left-tailed test, use the z value that corresponds to the area equivalent to α in Table E.

b. For a right-tailed test, use the z value that corresponds to the area equivalent to $1 - \alpha$.

c. For a two-tailed test, use the z value that corresponds to $\alpha/2$ for the left value. It will be negative. For the right value, use the z value that corresponds to the area equivalent to $1 - \alpha/2$. It will be positive.

EXAMPLE 8–2

Using Table E in Appendix A, find the critical value(s) for each situation and draw the appropriate figure, showing the critical region.

a. A left-tailed test with $\alpha = 0.10$.

b. A two-tailed test with $\alpha = 0.02$.

c. A right-tailed test with $\alpha = 0.005$.

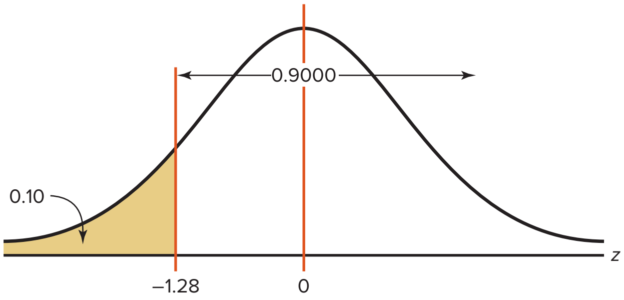
SOLUTION *a*

Step 1 Draw the figure and indicate the appropriate area. Since this is a left-tailed test, the area of 0.10 is located in the left tail, as shown in Figure 8–10.

Step 2 Find the area closest to 0.1000 in Table E. In this case, it is 0.1003. Find the z value that corresponds to the area 0.1003. It is −1.28. See Figure 8–10.

FIGURE 8–10

Critical Value and Critical Region for part *a* of Example 8–2



SOLUTION *b*

Step 1 Draw the figure and indicate the appropriate area. In this case, there are two areas equivalent to $\alpha/2$, or $0.02/2 = 0.01$.

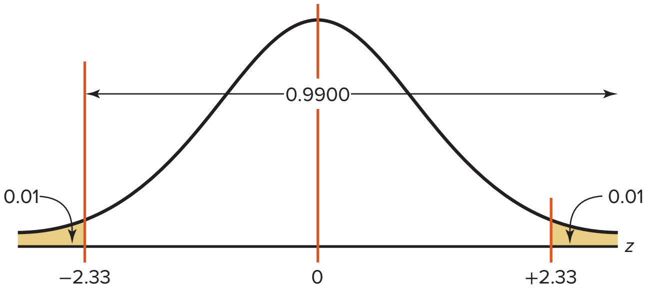
Step 2 For the left z critical value, find the area closest to $\alpha/2$, or $0.02/2 = 0.01$. In this case, it is 0.0099.

For the right z critical value, find the area closest to $1 - \alpha/2$, or $1 - 0.02/2 = 0.9900$. In this case, it is 0.9901.

Find the z values for each of the areas. For 0.0099, $z = -2.33$. For the area of 0.9901, $z = +2.33$. See Figure 8–11.

FIGURE 8–11

Critical Values and Critical Regions for Part *b* of Example 8–2

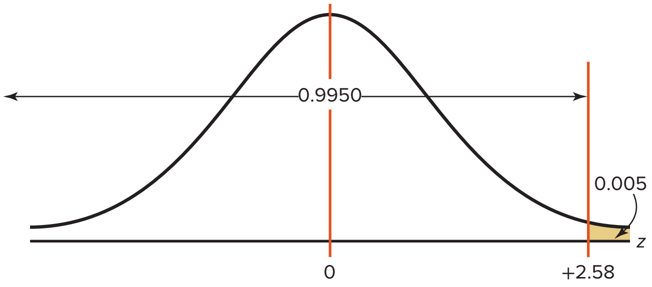


SOLUTION c

Step 1 Draw the figure and indicate the appropriate area. Since this is a right-tailed test, the area 0.005 is located in the right tail, as shown in Figure 8–12.

FIGURE 8–12

Critical Value and Critical Region for Part *c* of Example 8–2



Step 2 Find the area closest to $1 - \alpha$, or $1 - 0.005 = 0.9950$. In this case, it is 0.9949 or 0.9951.

The two z values corresponding to areas 0.9949 and 0.9951 are +2.57 and +2.58 respectively. Since 0.9500 is halfway between these two z values, find the average of the two z values $(+2.57 + 2.58) \div 2 = +2.575$. However, 2.58 is most often used. See Figure 8–12.

OBJECTIVE 4

State the five steps used in hypothesis testing.

In hypothesis testing, the following steps are recommended.

1. State the hypotheses. Be sure to state both the null and alternative hypotheses.
2. Design the study. This step includes selecting the correct statistical test, choosing a level of significance, and formulating a plan to carry out the study. The plan should include information such as the definition of the population, the way the sample will be selected, and the methods that will be used to collect the data.
3. Conduct the study and collect the data.
4. Evaluate the data. The data should be tabulated in this step, and the statistical test should be conducted. Finally, decide whether to reject or not reject the null hypothesis.
5. Summarize the results.

For the purposes of this chapter, a simplified version of the hypothesis-testing procedure will be used, since designing the study and collecting the data will be omitted. The steps are summarized in the Procedure Table.

Procedure Table

Solving Hypothesis-Testing Problems (Traditional Method)

Step 1 State the hypotheses and identify the claim.

Step 2 Find the critical value(s) from the appropriate table in Appendix A.

Step 3 Compute the test value.

Step 4 Make the decision to reject or not reject the null hypothesis.

Step 5 Summarize the results.

It is important to summarize the results of a statistical study correctly. The following table will help you when you summarize the results.

	Claim	
Decision	**Claim is H_0**	**Claim is H_1**
Reject H_0	There is enough evidence to reject the claim.	There is enough evidence to support the claim.
Do not reject H_0	There is not enough evidence to reject the claim.	There is not enough evidence to support the claim.

To further illustrate, the claim can be either the null or alternative hypothesis, and one should identify which it is. After the study is completed, the null hypothesis is either rejected or not rejected. From these two facts, the decision can be identified in the appropriate block.

For example, suppose a researcher claims that the mean weight of an adult animal of a particular species is 42 pounds. In this case, the claim would be the null hypothesis, H_0: $\mu = 42$, since the researcher is asserting that the parameter is a specific value. If the null hypothesis is rejected, the conclusion would be that there is enough evidence to reject the claim that the mean weight of the adult animal is 42 pounds.

On the other hand, suppose the researcher claims that the mean weight of the adult animals is not 42 pounds. The claim would be the alternative hypothesis H_1: $\mu \neq 42$. Furthermore, suppose that the null hypothesis is not rejected. The conclusion, then, would be that there is not enough evidence to support the claim that the mean weight of the adult animals is not 42 pounds.

Again, remember that nothing is being proved true or false. The statistician is only stating that there is or is not enough evidence to say that a claim is *probably* true or false. As noted previously, the only way to prove something would be to use the entire population under study, and usually this cannot be done, especially when the population is large.

Applying the Concepts **8–1**

Eggs and Your Health

The Incredible Edible Egg company recently found that eating eggs does not increase a person's blood serum cholesterol. Five hundred subjects participated in a study that lasted for 2 years. The participants were randomly assigned to either a no-egg group or a moderate-egg group. The blood serum cholesterol levels were checked at the beginning and at the end of the study. Overall, the groups' levels were not significantly different. The company reminds us that eating eggs is healthy if done in moderation. Many of the previous studies relating eggs and high blood serum cholesterol jumped to improper conclusions.

Using this information, answer these questions.

1. What prompted the study?
2. What is the population under study?
3. Was a sample collected?
4. What was the hypothesis?
5. Were data collected?
6. Were any statistical tests run?
7. What was the conclusion?

See page 485 for the answers.

Exercises 8–1

1. Define *null* and *alternative hypotheses,* and give an example of each.

2. What is meant by a type I error? A type II error? How are they related?

3. What is meant by a statistical test?

4. Explain the difference between a one-tailed and a two-tailed test.

5. What is meant by the critical region? The *noncritical* region?

6. What symbols are used to represent the null hypothesis and the alternative hypothesis?

7. What symbols are used to represent the probabilities of type I and type II errors?

8. Explain what is meant by a significant difference.

9. When should a one-tailed test be used? A two-tailed test?

10. In hypothesis testing, why can't the hypothesis be proved true?

11. Using the z table (Table E), find the critical value (or values) for each.

a. $\alpha = 0.10$, two-tailed test
b. $\alpha = 0.01$, right-tailed test
c. $\alpha = 0.005$, left-tailed test
d. $\alpha = 0.01$, left-tailed test
e. $\alpha = 0.05$, right-tailed test

12. Using the z table (Table E), find the critical value (or values) for each.

a. $\alpha = 0.02$, left-tailed test
b. $\alpha = 0.05$, right-tailed test
c. $\alpha = 0.01$, two-tailed test
d. $\alpha = 0.04$, left-tailed test
e. $\alpha = 0.02$, right-tailed test

13. For each conjecture, state the null and alternative hypotheses.

a. The average weight of dogs is 15.6 pounds.

b. The average distance a person lives away from a toxic waste site is greater than 10.8 miles.

c. The average farm size in 1970 was less than 390 acres.

d. The average number of miles a vehicle is driven per year is 12,603.

e. The average amount of money a person keeps in his or her checking account is less than \$24.

14. For each conjecture, state the null and alternative hypotheses.

a. The average age of first-year medical school students is at least 27 years.

b. The average experience (in seasons) for an NBA player is 4.71.

c. The average number of monthly visits/sessions on the Internet by a person at home has increased from 36 in 2009.

d. The average cost of a cell phone is \$79.95.

e. The average weight loss for a sample of people who exercise 30 minutes per day for 6 weeks is 8.2 pounds.

8–2 z Test for a Mean

OBJECTIVE ⑤

Test means when σ is known, using the z test.

In this chapter, two statistical tests will be explained: the z test is used when σ is known, and the t test is used when σ is unknown. This section explains the z test, and Section 8–3 explains the t test.

Many hypotheses are tested using a statistical test based on the following general formula:

$$\text{Test value} = \frac{(\text{observed value}) - (\text{expected value})}{\text{standard error}}$$

The observed value is the statistic (such as the sample mean) that is computed from the sample data. The expected value is the parameter (such as the population mean) that you would expect to obtain if the null hypothesis were true—in other words, the hypothesized value. The denominator is the standard error of the statistic being tested (in this case, the standard error of the mean).

The *z* test is defined formally as follows.

The ***z* test** is a statistical test for the mean of a population. It can be used either when $n \geq 30$ or when the population is normally distributed and σ is known.

The formula for the *z* test is

$$z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

where $\overline{X}$ = sample mean
μ = hypothesized population mean
σ = population standard deviation
n = sample size

For the *z* test, the observed value is the value of the sample mean. The expected value is the value of the population mean, assuming that the null hypothesis is true. The denominator $\sigma/\sqrt{n}$ is the standard error of the mean.

The formula for the *z* test is the same formula shown in Chapter 6 for the situation where you are using a distribution of sample means. Recall that the central limit theorem allows you to use the standard normal distribution to approximate the distribution of sample means when $n \geq 30$.

The assumptions for the *z* test when σ is known are given next.

Assumptions for the *z* Test for a Mean When σ Is Known

1. The sample is a random sample.
2. Either $n \geq 30$ or the population is normally distributed when $n < 30$.

In this text, the assumptions will be stated in the exercises; however, when encountering statistics in other situations, you must check to see that these assumptions have been met before proceeding.

As stated in Section 8–1, there are five steps for solving *hypothesis-testing* problems:

Step 1 State the hypotheses and identify the claim.

Step 2 Find the critical value(s).

Step 3 Compute the test value.

Step 4 Make the decision to reject or not reject the null hypothesis.

Step 5 Summarize the results.

Note: Your first encounter with hypothesis testing can be somewhat challenging and confusing, since there are many new concepts being introduced at the same time. *To understand all the concepts, you must carefully follow each step in the examples and try each exercise that is assigned.* Only after careful study and patience will these concepts become clear.

Again, *z* values are rounded to two decimal places because they are given to two decimal places in Table E.

EXAMPLE 8–3 Ages of Medical Doctors

A researcher believes that the mean age of medical doctors in a large hospital system is older than the average age of doctors in the United States, which is 46. Assume the population standard deviation is 4.2 years. A random sample of 30 doctors from the system is selected, and the mean age of the sample is 48.6. Test the claim at $\alpha = 0.05$.

Source: American Averages.

SOLUTION

Step 1 State the hypotheses and identify the claim.

H_0: $\mu = 46$ and H_1: $\mu > 46$ (claim)

Step 2 Find the critical value. Since $\alpha = 0.05$ and the test is right-tailed, the critical value is $z = +1.65$.

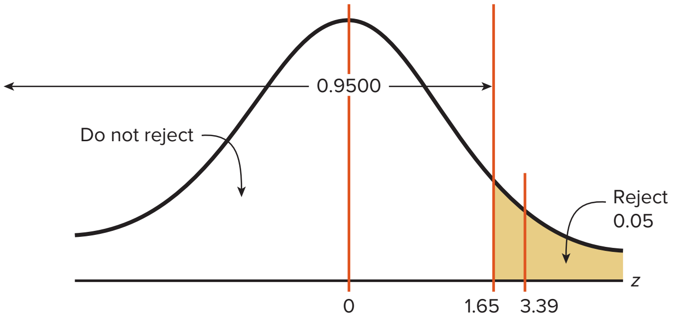
Step 3 Compute the test value.

$$z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{48.6 - 46}{4.2/\sqrt{30}} = 3.39$$

Step 4 Make the decision. Since 3.39 is greater than the critical value 1.65, the decision is to reject the null hypothesis. This is shown in Figure 8–13.

FIGURE 8–13

Summary of the *z* test of Example 8–3



Step 5 Summarize the results. There is enough evidence to support the claim that the average age of the doctors in the system is greater than the national average of 46.

Comment: In Example 8–3, the difference is said to be statistically significant. However, when the null hypothesis is rejected, there is always a chance of a type I error. In this case, the probability of a type I error is at most 0.05, or 5%.

EXAMPLE 8–4 SAT Tests

For a specific year, the average score on the SAT Math test was 515.* The variable is normally distributed, and the population standard deviation is 100. The same superintendent in the previous example wishes to see if her students scored significantly below the national average on the test. She randomly selected 36 student scores, as shown. At $\alpha = 0.10$, is there enough evidence to support the claim?

**Source:* www.chacha.com

496	506	507	505	438	499
505	522	531	762	513	493
522	668	543	519	349	506
519	516	714	517	511	551
287	523	576	516	515	500
243	509	523	503	414	504

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon \mu = 515 \qquad \text{and} \qquad H_1\colon \mu < 515 \text{ (claim)}$$

Step 2 Find the critical value. Since $\alpha = 0.10$ and the test is a left-tailed test, the critical value is -1.28.

Step 3 Compute the test value. Since the exercise gives raw data, it is necessary to find the mean of the data.

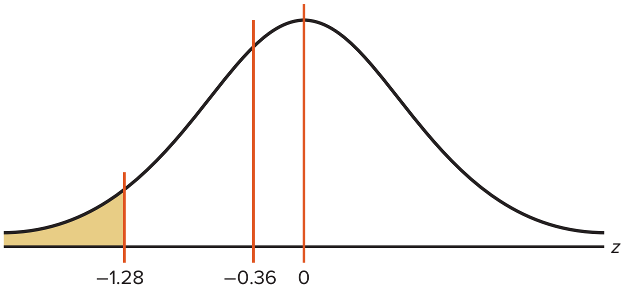
$$\overline{X} = \frac{496 + 505 + 522 + \cdots + 504}{36} = \frac{18{,}325}{36} = 509.028$$

$$z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{509.028 - 515}{100/\sqrt{36}} = -0.36$$

Step 4 Make the decision. Since the test value, −0.36, falls in the noncritical region, the decision is to not reject the null hypothesis. See Figure 8–14.

FIGURE 8–14

Critical and Test Values for Example 8–4



Step 5 Summarize the results. There is not enough evidence to support the claim that the students scored below the national average.

Comment: Even though in Example 8–4 the sample mean of 509.028 is lower than the hypothesized population mean of 515, it is not *significantly* lower. Hence, the difference may be due to chance. When the null hypothesis is not rejected, there is still a probability of a type II error, i.e., of not rejecting the null hypothesis when it is false.

The probability of a type II error is not easily ascertained. Further explanation about the type II error is given in Section 8–6. For now, it is only necessary to realize that the probability of type II error exists when the decision is to not reject the null hypothesis.

Also note that when the null hypothesis is not rejected, it cannot be accepted as true. There is merely not enough evidence to say that it is false. This guideline may sound a little confusing, but the situation is analogous to a jury trial. The verdict is either guilty or not guilty and is based on the evidence presented. If a person is judged not guilty, it does not mean that the person is proved innocent; it only means that there was not enough evidence to reach the guilty verdict.

EXAMPLE 8–5 Cost of Rehabilitation

The Medical Rehabilitation Education Foundation reports that the average cost of rehabilitation for stroke victims is $24,672. To see if the average cost of rehabilitation is different at a particular hospital, a researcher selects a random sample of 35 stroke victims at the hospital and finds that the average cost of their rehabilitation is $26,343. The standard deviation of the population is $3251. At $\alpha = 0.01$, can it be concluded that the average cost of stroke rehabilitation at a particular hospital is different from $24,672?

Source: Snapshot, *USA TODAY.*

SOLUTION

Step 1 State the hypotheses and identify the claim.

H_0: $\mu = \$24,672$ and H_1: $\mu \neq \$24,672$ (claim)

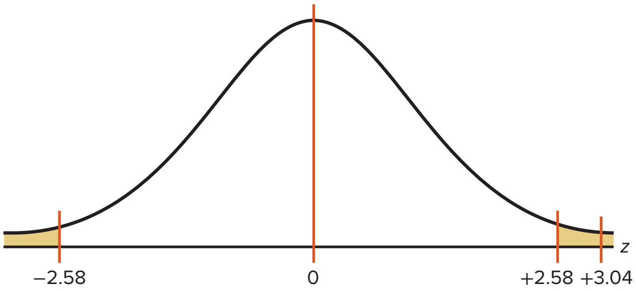
Step 2 Find the critical values. Since $\alpha = 0.01$ and the test is a two-tailed test, the critical values are +2.58 and −2.58.

Step 3 Compute the test value.

$$z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{26,343 - 24,672}{3251/\sqrt{35}} = 3.04$$

Step 4 Make the decision. Reject the null hypothesis, since the test value falls in the critical region, as shown in Figure 8–15.

FIGURE 8–15
Critical and Test Values for Example 8–5



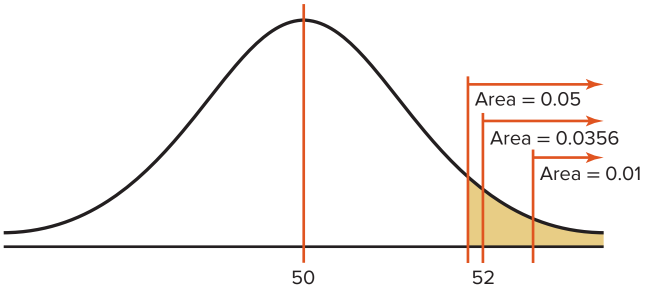
Step 5 Summarize the results. There is enough evidence to support the claim that the average cost of rehabilitation at the particular hospital is different from $24,672.

P-Value Method for Hypothesis Testing

Statisticians usually test hypotheses at the common α levels of 0.05 or 0.01 and sometimes at 0.10. Recall that the choice of the level depends on the seriousness of the type I error. Besides listing an α value, many computer statistical packages give a *P*-value for hypothesis tests.

The ***P*-value** (or probability value) is the probability of getting a sample statistic (such as the sample mean) or a more extreme sample statistic in the direction of the alternative hypothesis when the null hypothesis is true.

FIGURE 8–16
Comparison of α Values and P-Values



In other words, the P-value is the actual area under the standard normal distribution curve (or other curve, depending on what statistical test is being used) representing the probability of a particular sample statistic or a more extreme sample statistic occurring if the null hypothesis is true.

For example, suppose that an alternative hypothesis is H_1: $\mu > 50$ and the mean of a sample is $\overline{X} = 52$. If the computer printed a P-value of 0.0356 for a statistical test, then the probability of getting a sample mean of 52 or greater is 0.0356 if the true population mean is 50 (for the given sample size and standard deviation). The relationship between the P-value and the α value can be explained in this manner. For $P = 0.0356$, the null hypothesis would be rejected at $\alpha = 0.05$ but not at $\alpha = 0.01$. See Figure 8–16.

When the hypothesis test is two-tailed, the area in one tail must be doubled. For a two-tailed test, if α is 0.05 and the area in one tail is 0.0356, the P-value will be $2(0.0356) = 0.0712$. That is, the null hypothesis should not be rejected at $\alpha = 0.05$, since 0.0712 is greater than 0.05. In summary, then, if the P-value is less than α, reject the null hypothesis. If the P-value is greater than α, do not reject the null hypothesis.

Decision Rule When Using a *P*-Value

If P-value $\leq \alpha$, reject the null hypothesis.
If P-value $> \alpha$, do not reject the null hypothesis.

The P-values for the z test can be found by using Table E in Appendix A. First find the area under the standard normal distribution curve corresponding to the z test value. For a left-tailed test, use the area given in the table; for a right-tailed test, use 1.0000 minus the area given in the table. To get the P-value for a two-tailed test, double the area you found in the tail. This procedure is shown in step 3 of Examples 8–6 and 8–7.

The P-value method for testing hypotheses differs from the traditional method somewhat. The steps for the P-value method are summarized next.

Procedure Table

Solving Hypothesis-Testing Problems (*P*-Value Method)

Step 1 State the hypotheses and identify the claim.

Step 2 Compute the test value.

Step 3 Find the P-value.

Step 4 Make the decision.

Step 5 Summarize the results.

Examples 8–6 and 8–7 show how to use the *P*-value method to test hypotheses.

EXAMPLE 8–6 Cost of College Tuition

A researcher wishes to test the claim that the average cost of tuition and fees at a four-year public college is greater than \$5700. She selects a random sample of 36 four-year public colleges and finds the mean to be \$5950. The population standard deviation is \$659. Is there evidence to support the claim at $\alpha = 0.05$? Use the *P*-value method.

Source: Based on information from the College Board.

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon \mu = \$5700 \qquad \text{and} \qquad H_1\colon \mu > \$5700 \text{ (claim)}$$

Step 2 Compute the test value.

$$z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{5950 - 5700}{659/\sqrt{36}} = 2.28$$

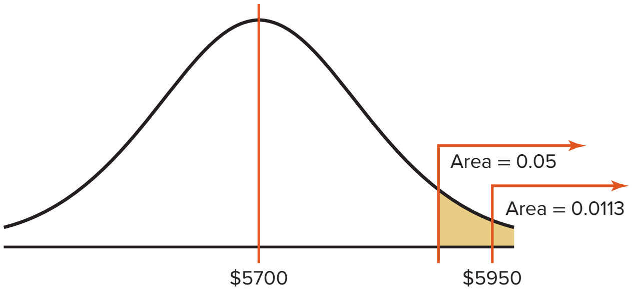
Step 3 Find the *P*-value. Using Table E in Appendix A, find the corresponding area under the normal distribution for $z = 2.28$. It is 0.9887. Subtract this value for the area from 1.0000 to find the area in the right tail.

$$1.0000 - 0.9887 = 0.0113$$

Hence, the *P*-value is 0.0113.

Step 4 Make the decision. Since the *P*-value is less than 0.05, the decision is to reject the null hypothesis. See Figure 8–17.

FIGURE 8–17
P-Value and α Value for Example 8–6



Step 5 Summarize the results. There is enough evidence to support the claim that the tuition and fees at four-year public colleges are greater than \$5700.

Note: Had the researcher chosen $\alpha = 0.01$, the null hypothesis would not have been rejected since the *P*-value (0.0113) is greater than 0.01.

EXAMPLE 8–7 Wind Speed

A researcher claims that the average wind speed in a certain city is 8 miles per hour. A sample of 32 days has an average wind speed of 8.2 miles per hour. The standard deviation of the population is 0.6 mile per hour. At $\alpha = 0.05$, is there enough evidence to reject the claim? Use the *P*-value method.

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon \mu = 8 \text{ (claim)} \qquad \text{and} \qquad H_1\colon \mu \neq 8$$

Step 2 Compute the test value.

$$z = \frac{8.2 - 8}{0.6/\sqrt{32}} = 1.89$$

Step 3 Find the *P*-value. Using Table E, find the corresponding area for $z = 1.89$. It is 0.9706. Subtract the value from 1.0000.

$$1.0000 - 0.9706 = 0.0294$$

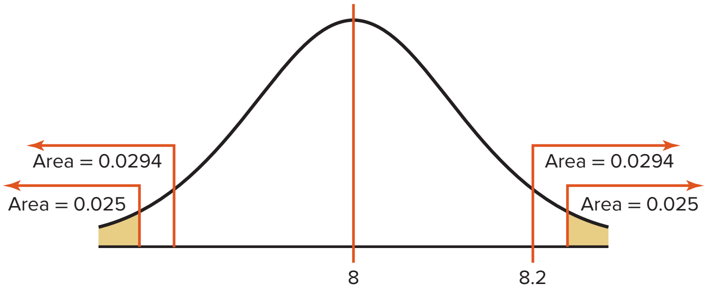
Since this is a two-tailed test, the area of 0.0294 must be doubled to get the *P*-value.

$$2(0.0294) = 0.0588$$

Step 4 Make the decision. The decision is to not reject the null hypothesis, since the *P*-value is greater than 0.05. See Figure 8–18.

FIGURE 8–18

P-Values and α Values for Example 8–7



Step 5 Summarize the results. There is not enough evidence to reject the claim that the average wind speed is 8 miles per hour.

In Examples 8–6 and 8–7, the *P*-value and the α value were shown on a normal distribution curve to illustrate the relationship between the two values; however, it is not necessary to draw the normal distribution curve to make the decision whether to reject the null hypothesis. You can use the following rule:

Decision Rule When Using a *P*-Value

If *P*-value $\leq \alpha$, reject the null hypothesis.
If *P*-value $> \alpha$, do not reject the null hypothesis.

In Example 8–6, *P*-value = 0.0113 and $\alpha = 0.05$. Since *P*-value $\leq \alpha$, the null hypothesis was rejected. In Example 8–7, *P*-value = 0.0588 and $\alpha = 0.05$. Since *P*-value $> \alpha$, the null hypothesis was not rejected.

The *P*-values given on calculators and computers are slightly different from those found with Table E. This is so because z values and the values in Table E have been rounded. *Also, most calculators and computers give the exact P-value for two-tailed tests, so it should not be doubled (as it should when the area found in Table E is used).*

A clear distinction between the α value and the *P*-value should be made. The α value is chosen by the researcher *before* the statistical test is conducted. The *P*-value is computed after the sample mean has been found.

There are two schools of thought on *P*-values. Some researchers do not choose an α value but report the *P*-value and allow the reader to decide whether the null hypothesis should be rejected.

In this case, the following guidelines can be used, but be advised that these guidelines are not written in stone, and some statisticians may have other opinions.

Guidelines for *P*-Values

If *P*-value ≤ 0.01, reject the null hypothesis. The difference is highly significant.

If *P*-value > 0.01 but *P*-value ≤ 0.05, reject the null hypothesis. The difference is significant.

If *P*-value > 0.05 but *P*-value ≤ 0.10, consider the consequences of type I error before rejecting the null hypothesis.

If *P*-value > 0.10, do not reject the null hypothesis. The difference is not significant.

Others decide on the α value in advance and use the *P*-value to make the decision, as shown in Examples 8–6 and 8–7. A note of caution is needed here: If a researcher selects $\alpha = 0.01$ and the *P*-value is 0.03, the researcher may decide to change the α value from 0.01 to 0.05 so that the null hypothesis will be rejected. This, of course, should not be done. If the α level is selected in advance, it should be used in making the decision.

One additional note on hypothesis testing is that the researcher should distinguish between *statistical significance* and *practical significance*. When the null hypothesis is rejected at a specific significance level, it can be concluded that the difference is probably not due to chance and thus is statistically significant. However, the results may not have any practical significance. For example, suppose that a new fuel additive increases the miles per gallon that a car can get by $\frac{1}{4}$ mile for a sample of 1000 automobiles. The results may be statistically significant at the 0.05 level, but it would hardly be worthwhile to market the product for such a small increase. Hence, there is no practical significance to the results. It is up to the researcher to use common sense when interpreting the results of a statistical test.

Applying the Concepts **8–2**

Car Thefts

You recently received a job with a company that manufactures an automobile antitheft device. To conduct an advertising campaign for the product, you need to make a claim about the number of automobile thefts per year. Since the population of various cities in the United States varies, you decide to use rates per 10,000 people. (The rates are based on the number of people living in the cities.) Your boss said that last year the theft rate per 10,000 people was 44 vehicles. You want to see if it has changed. The following are rates per 10,000 people for 36 randomly selected locations in the United States. Assume $\sigma = 30.3$.

55	42	125	62	134	73
39	69	23	94	73	24
51	55	26	66	41	67
15	53	56	91	20	78
70	25	62	115	17	36
58	56	33	75	20	16

Source: Based on information from the National Insurance Crime Bureau.

Using this information, answer these questions.

1. What hypotheses would you use?
2. Is the sample considered small or large?
3. What assumption must be met before the hypothesis test can be conducted?
4. Which probability distribution would you use?
5. Would you select a one- or two-tailed test? Why?
6. What critical value(s) would you use?

7. Conduct a hypothesis test.
8. What is your decision?
9. What is your conclusion?
10. Write a brief statement summarizing your conclusion.
11. If you lived in a city whose population was about 50,000, how many automobile thefts per year would you expect to occur?

See pages 485–486 for the answers.

Exercises 8–2

For Exercises 1 through 25, perform each of the following steps.

- *a.* State the hypotheses and identify the claim.
- *b.* Find the critical value(s).
- *c.* Compute the test value.
- *d.* Make the decision.
- *e.* Summarize the results.

Use diagrams to show the critical region (or regions), and use the traditional method of hypothesis testing unless otherwise specified.

1. **Warming and Ice Melt** The average depth of the Hudson Bay is 305 feet. Climatologists were interested in seeing if warming and ice melt were affecting the water level. Fifty-five measurements over a period of randomly selected weeks yielded a sample mean of 306.2 feet. The population variance is known to be 3.6. Can it be concluded at the 0.05 level of significance that the average depth has increased? Is there evidence of what caused this to happen?

   Source: World Almanac and Book of Facts 2010.

2. **Facebook Friends** Many people believe that the average number of Facebook friends is 338. The population standard deviation is 43.2. A random sample of 50 high school students in a particular county revealed that the average number of Facebook friends was 350. At $\alpha = 0.05$, is there sufficient evidence to conclude that the mean number of friends is greater than 338?

3. **Revenue of Large Businesses** A researcher estimates that the average revenue of the largest businesses in the United States is greater than \$24 billion. A random sample of 50 companies is selected, and the revenues (in billions of dollars) are shown. At $\alpha = 0.05$, is there enough evidence to support the researcher's claim? Assume $\sigma = 28.7$.

178	122	91	44	35
61	56	46	20	32
30	28	28	20	27
29	16	16	19	15
41	38	36	15	25
31	30	19	19	19
24	16	15	15	19
25	25	18	14	15
24	23	17	17	22
22	21	20	17	20

Source: New York Times Almanac.

4. **Moviegoers** The average "moviegoer" sees 8.5 movies a year. A *moviegoer* is defined as a person who sees at least one movie in a theater in a 12-month period. A random sample of 40 moviegoers from a large university revealed that the average number of movies seen per person was 9.6. The population standard deviation is 3.2 movies. At the 0.05 level of significance, can it be concluded that this represents a difference from the national average?

   Source: MPAA Study.

5. **Sick Days in Bed** A researcher wishes to see if the average number of sick days a worker takes per year is greater than 5. A random sample of 32 workers at a large department store had a mean of 5.6. The standard deviation of the population is 1.2. Is there enough evidence to support the researcher's claim at $\alpha = 0.01$?

6. **Cost of Building a Home** According to the National Association of Home Builders, the average cost of building a home in the Northeast is \$117.91 per square foot. A random sample of 36 new homes indicated that the mean cost was \$122.57 and the standard deviation was \$20. Can it be concluded that the mean cost differs from \$117.91, using the 0.10 level of significance?

7. **Heights of 1-Year-Olds** The average 1-year-old (both genders) is 29 inches tall. A random sample of 30 1-year-olds in a large day care franchise resulted in

the following heights. At $\alpha = 0.05$, can it be concluded that the average height differs from 29 inches? Assume $\sigma = 2.61$.

25	32	35	25	30	26.5	26	25.5	29.5	32
30	28.5	30	32	28	31.5	29	29.5	30	34
29	32	27	28	33	28	27	32	29	29.5

Source: www.healthepic.com

8. **Salaries of Government Employees** The mean salary of federal government employees on the General Schedule is \$59,593. The average salary of 30 randomly selected state employees who do similar work is \$58,800 with $\sigma = \$1500$. At the 0.01 level of significance, can it be concluded that state employees earn on average less than federal employees?

Source: New York Times Almanac.

9. **Telephone Calls** A researcher knew that before cell phones, a person made on average 2.8 calls per day. He believes that the number of calls made per day today is higher. He selects a random sample of 30 individuals who use a cell phone and asks them to keep track of the number of calls that they made on a certain day. The mean was 3.1. At $\alpha = 0.01$ is there enough evidence to support the researcher's claim? The standard deviation for the population found by a previous study is 0.8. Would the null hypothesis be rejected at $\alpha = 0.05$?

10. **Daily Driving** The average number of miles a person drives per day is 24. A researcher wishes to see if people over age 60 drive less than 24 miles per day. She selects a random sample of 49 drivers over the age of 60 and finds that the mean number of miles driven is 22.8. The population standard deviation is 3.5 miles. At $\alpha = 0.05$ is there sufficient evidence that those drivers over 60 years old drive less on average than 24 miles per day?

11. **Weight Loss of Newborns** An obstetrician read that a newborn baby loses on average 7 ounces in the first two days of his or her life. He feels that in the hospital where he works, the average weight loss of a newborn baby is less than 7 ounces. A random sample of 32 newborn babies has a mean weight loss of 6.5 ounces. The population standard deviation is 1.8 ounces. Is there enough evidence at $\alpha = 0.01$ to support his claim?

12. **Student Expenditures** The average expenditure per student (based on average daily attendance) for a certain school year was \$10,337 with a population standard deviation of \$1560. A survey for the next school year of 150 randomly selected students resulted in a sample mean of \$10,798. Do these results indicate that the average expenditure has changed? Choose your own level of significance.

Source: World Almanac.

13. **Dress Shirts** In a previous study conducted several years ago, a man owned on average 15 dress shirts. The standard deviation of the population is 3. A researcher wishes to see if that average has changed. He selected a random sample of 42 men and found that the average number of dress shirts that they owned was 13.8. At $\alpha = 0.05$, is there enough evidence to support the claim that the average has changed?

14. **Prison Sentences** The average length of prison term in the United States for white collar crime is 34.9 months. At $\alpha = 0.04$, is there sufficient evidence to conclude that the average stay differs from 34.9 months?

Source: californiawatch.org

15. **Reject or Not** State whether the null hypothesis should be rejected on the basis of the given *P*-value.

    a. *P*-value = 0.258, $\alpha = 0.05$, one-tailed test
    b. *P*-value = 0.0684, $\alpha = 0.10$, two-tailed test
    c. *P*-value = 0.0153, $\alpha = 0.01$, one-tailed test
    d. *P*-value = 0.0232, $\alpha = 0.05$, two-tailed test
    e. *P*-value = 0.002, $\alpha = 0.01$, one-tailed test

16. **Soft Drink Consumption** A researcher claims that the yearly consumption of soft drinks per person is 52 gallons. In a sample of 50 randomly selected people, the mean of the yearly consumption was 56.3 gallons. The standard deviation of the population is 3.5 gallons. Find the *P*-value for the test. On the basis of the *P*-value, is the researcher's claim valid?

Source: U.S. Department of Agriculture.

17. **Stopping Distances** A study found that the average stopping distance of a school bus traveling 50 miles per hour was 264 feet. A group of automotive engineers decided to conduct a study of its school buses and found that for 20 randomly selected buses, the average stopping distance of buses traveling 50 miles per hour was 262.3 feet. The standard deviation of the population was 3 feet. Test the claim that the average stopping distance of the company's buses is actually less than 264 feet. Find the *P*-value. On the basis of the *P*-value, should the null hypothesis be rejected at $\alpha = 0.01$? Assume that the variable is normally distributed.

Source: Snapshot, *USA TODAY.*

18. **Copy Machine Use** A store manager hypothesizes that the average number of pages a person copies on the store's copy machine is less than 40. A random sample of 50 customers' orders is selected. At $\alpha = 0.01$, is there enough evidence to support the claim? Use the *P*-value hypothesis-testing method. Assume $\sigma = 30.9$.

2	2	2	5	32
5	29	8	2	49
21	1	24	72	70
21	85	61	8	42
3	15	27	113	36
37	5	3	58	82
9	2	1	6	9
80	9	51	2	122
21	49	36	43	61
3	17	17	4	1

19. Medical School Applications A medical college dean read that the average number of applications a potential medical school student sends is 7.8. She thinks that the mean is higher. So she selects a random sample of 35 applicants and asks each how many medical schools they applied to. The mean of the sample is 8.7. The population standard deviation is 2.6. Test her claim at $\alpha = 0.01$.

20. Breaking Strength of Cable A special cable has a breaking strength of 800 pounds. The standard deviation of the population is 12 pounds. A researcher selects a random sample of 20 cables and finds that the average breaking strength is 793 pounds. Can he reject the claim that the breaking strength is 800 pounds? Find the *P*-value. Should the null hypothesis be rejected at $\alpha = 0.01$? Assume that the variable is normally distributed.

21. Farm Sizes The average farm size in the United States is 444 acres. A random sample of 40 farms in Oregon indicated a mean size of 430 acres, and the population standard deviation is 52 acres. At $\alpha = 0.05$, can it be concluded that the average farm in Oregon differs from the national mean? Use the *P*-value method.

Source: New York Times Almanac.

22. Farm Sizes Ten years ago, the average acreage of farms in a certain geographic region was 65 acres. The standard deviation of the population was 7 acres. A recent study consisting of 22 randomly selected farms showed that the average was 63.2 acres per farm. Test the claim, at $\alpha = 0.10$, that the average has not changed by finding the *P*-value for the test. Assume that σ has not changed and the variable is normally distributed.

23. Transmission Service A car dealer recommends that transmissions be serviced at 30,000 miles. To see whether her customers are adhering to this recommendation, the dealer selects a random sample of 40 customers and finds that the average mileage of the automobiles serviced is 30,456. The standard deviation of the population is 1684 miles. By finding the *P*-value, determine whether the owners are having their transmissions serviced at 30,000 miles. Use $\alpha = 0.10$. Do you think the α value of 0.10 is an appropriate significance level?

24. Speeding Tickets A motorist claims that the South Boro Police issue an average of 60 speeding tickets per day. These data show the number of speeding tickets issued each day for a randomly selected period of 30 days. Assume σ is 13.42. Is there enough evidence to reject the motorist's claim at $\alpha = 0.05$? Use the *P*-value method.

72	45	36	68	69	71	57	60
83	26	60	72	58	87	48	59
60	56	64	68	42	57	57	
58	63	49	73	75	42	63	

25. Sick Days A manager states that in his factory, the average number of days per year missed by the employees due to illness is less than the national average of 10. The following data show the number of days missed by 40 randomly selected employees last year. Is there sufficient evidence to believe the manager's statement at $\alpha = 0.05$? $\sigma = 3.63$. Use the *P*-value method.

0	6	12	3	3	5	4	1
3	9	6	0	7	6	3	4
7	4	7	1	0	8	12	3
2	5	10	5	15	3	2	5
3	11	8	2	2	4	1	9

Extending the Concepts

26. Significance Levels Suppose a statistician chose to test a hypothesis at $\alpha = 0.01$. The critical value for a right-tailed test is +2.33. If the test value were 1.97, what would the decision be? What would happen if, after seeing the test value, she decided to choose $\alpha = 0.05$? What would the decision be? Explain the contradiction, if there is one.

27. Hourly Wage The president of a company states that the average hourly wage of her employees is \$8.65. A random sample of 50 employees has the distribution shown. At $\alpha = 0.05$, is the president's statement believable? Assume $\sigma = 0.105$.

Class	Frequency
8.35–8.43	2
8.44–8.52	6
8.53–8.61	12
8.62–8.70	18
8.71–8.79	10
8.80–8.88	2

Technology Step by Step

TI-84 Plus Step by Step

Hypothesis Test for the Mean and the *z* Distribution (Data)

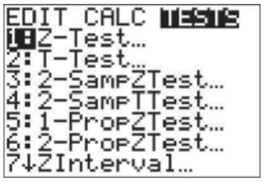
1. Enter the data values into L_1.
2. Press **STAT** and move the cursor to TESTS.
3. Press **1** for ZTest.

4. Move the cursor to Data and press **ENTER.**
5. Type in the appropriate values.
6. Move the cursor to the appropriate alternative hypothesis and press **ENTER.**
7. Move the cursor to Calculate and press **ENTER.**

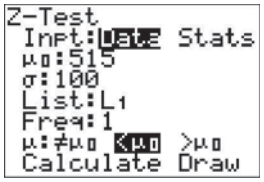
Example TI8–1

This relates to Example 8–4 from the text. At the 10% significance level, test the claim that $\mu < 515$, given the data values.

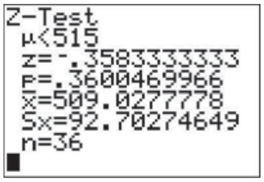
496	506	507	505	438	499	505	522	531	762	513	493
522	668	543	519	349	506	519	516	714	517	511	551
287	523	576	516	515	500	243	509	523	503	414	504



```
EDIT CALC TESTS
1:Z-Test…
2:T-Test…
3:2-SampZTest…
4:2-SampTTest…
5:1-PropZTest…
6:2-PropZTest…
7↓ZInterval…
```



```
Z-Test
 Inpt:Data Stats
 μ0:515
 σ:100
 List:L1
 Freq:1
 μ:≠μ0 <μ0 >μ0
 Calculate Draw
```



```
Z-Test
 μ<515
 z=-.3583333333
 p=.3600469966
 x̄=509.0277778
 Sx=92.70274649
 n=36
```

The test statistic is $z = -0.3583333333$, and the P-value is 0.3600469966.

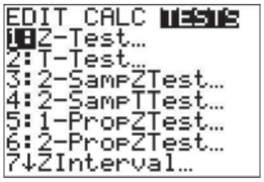
Hypothesis Test for the Mean and the *z* Distribution (Statistics)

1. Press **STAT** and move the cursor to TESTS.
2. Press **1** for ZTest.
3. Move the cursor to Stats and press **ENTER.**
4. Type in the appropriate values.
5. Move the cursor to the appropriate alternative hypothesis and press **ENTER.**
6. Move the cursor to Calculate and press **ENTER.**

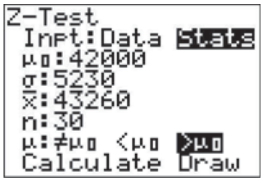
Example TI8–2

At the 5% significance level, test the claim that $\mu > 42{,}000$ given $\sigma = 5230$, $\overline{X} = 43{,}260$, and $n = 30$.

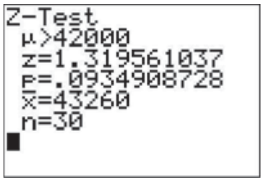
The test statistic is $z = 1.319561037$, and the P-value is 0.0934908728.



```
EDIT CALC TESTS
1:Z-Test…
2:T-Test…
3:2-SampZTest…
4:2-SampTTest…
5:1-PropZTest…
6:2-PropZTest…
7↓ZInterval…
```



```
Z-Test
 Inpt:Data Stats
 μ0:42000
 σ:5230
 x̄:43260
 n:30
 μ:≠μ0 <μ0 >μ0
 Calculate Draw
```



```
Z-Test
 μ>42000
 z=1.319561037
 p=.0934908728
 x̄=43260
 n=30
```

EXCEL
Step by Step

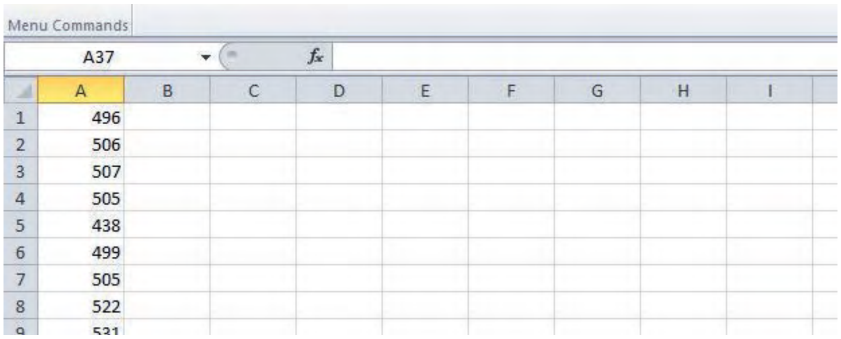
Hypothesis Test for the Mean: *z* Test

Excel does not have a procedure to conduct a hypothesis test for a single population mean. However, you may conduct the test of the mean by using the MegaStat Add-in available in your online resources. If you have not installed this add-in, do so, following the instructions from the Chapter 1 Excel Step by Step.

Example XL8–1

This example relates to Example 8–4 from the text. At the 10% level of significance, test the claim that $\mu < 515$. The MegaStat z test uses the P-value method. Therefore, it is not necessary to enter a significance level.

1. Enter the data from Example 8–4 into column A of a new Excel worksheet.
2. From the Toolbar, select Add-Ins, **MegaStat>Hypothesis Tests>Mean vs. Hypothesized Value.** *Note:* You may need to open MegaStat from the MegaStat.xls file on your computer's hard drive.



	A
1	496
2	506
3	507
4	505
5	438
6	499
7	505
8	522
9	531

3. Select data input and type **A1:A36** as the Input Range.
4. Type **515** for the Hypothesized mean and select the Alternative "less than."
5. Select *z* test and click [OK].

The result of the procedure is shown next.

```
Hypothesis Test: Mean vs. Hypothesized Value
515.000    Hypothesized value
509.028    Mean data
 92.703    Std. dev.
 15.450    Std. error
     36    n

  −0.39    z
 0.3495    P-value (one-tailed, lower)
```

MINITAB
Step by Step

Hypothesis Test for the Mean: *z* Test

MINITAB can be used to calculate the test statistic and the *P*-value. Although the *P*-value approach does not require a critical value from the table, MINITAB can also calculate a critical value as in Example 8–2.

Example 8–2

Find the critical value(s) for each situation and draw the appropriate figure.

Step 1 To find the critical value of *z* for a left-tailed test with $\alpha = 0.10$, select **Graph>Probability Distribution Plot,** then **View Probability,** and then click **[OK].**

Step 2 The Distribution should be set for Normal with a mean of 0 and a standard deviation of 1.

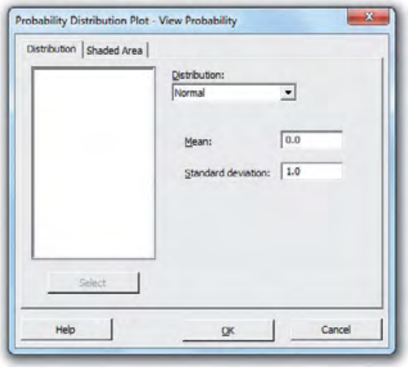
Step 3 Click the tab for **Shaded Area.**

a) Select **Left Tail.**

b) Select the ratio button for **Probability.**

c) Type in the value of alpha for probability, **0.10.**

d) Click **[OK].** The critical value of *z* to three decimal places is −1.282.

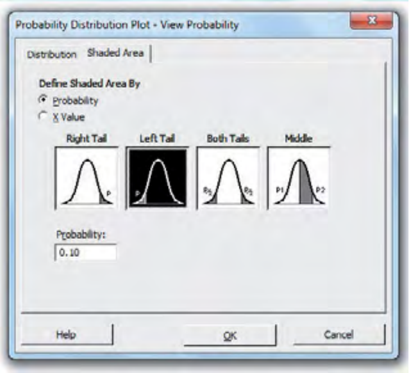


Probability Distribution Plot - View Probability

Distribution: Normal

Mean: 0.0

Standard deviation: 1.0



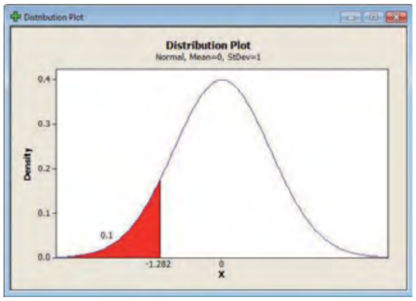
Probability Distribution Plot - View Probability

Define Shaded Area By

● Probability

○ X Value

Probability: 0.10



You may click the Edit Last Dialog button and then change the settings for additional critical values.

Example 8–3 Intelligence Tests

MINITAB will calculate the test statistic and *P*-value from the summary statistics. There are no data for this example.

Step 1 Select **Stat>Basic Statistics> 1-sample z.** This is the same menu item used to obtain a *z* interval in Chapter 7.

Step 2 Select **Summarized Data** from the drop down menu.

a) Type in the sample size of **30.**

b) Type in the sample mean of **106.4.**

c) Type in the population standard deviation of **15.**

Step 3 Check the box for Perform hypothesis test, then type in the Hypothesized value of **101.5.**

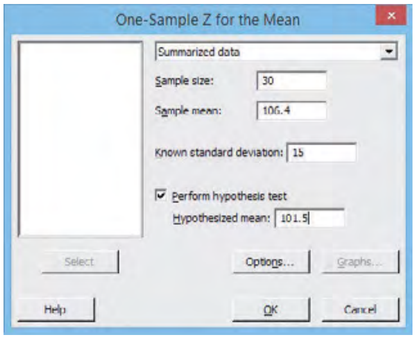
Step 4 Click the button for **[Options]**.

a) Type the percentage for the Confidence level, **95.**

b) Click the drop-down menu for the Alternative hypothesis, "Mean > hypothesized mean."

Step 5 Click **[OK]** twice.

In the session window you will see the results including a confidence interval estimate. The test statistic $z = 1.79$ with a *P*-value of 0.037. Since the *P*-value is smaller than α, the null hypothesis is rejected.



One-Sample Z for the Mean

Summarized data

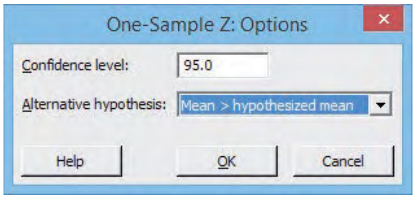
Sample size: 30

Sample mean: 106.4

Known standard deviation: 15

☑ Perform hypothesis test

Hypothesized mean: 101.5



One-Sample Z: Options

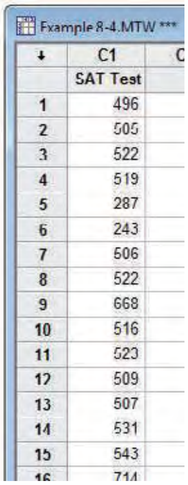
Confidence level: 95.0

Alternative hypothesis: Mean > hypothesized mean

In the session window you will see the results.

```
One-Sample z

Test of μ = 101.5 vs > 101.5
The assumed standard deviation = 15

                            95% Lower
 N     Mean    SE Mean         Bound      Z       P
30   106.40       2.74        101.90   1.79   0.037
```

Example 8–4 SAT Tests



	C1
	SAT Test
1	496
2	505
3	522
4	519
5	287
6	243
7	606
8	522
9	668
10	516
11	523
12	509
13	507
14	531
15	543

MINITAB will calculate the test statistic and *P*-value.

Step 1 Type the data into a new MINITAB worksheet. All 36 values must be in C1. The label SAT Test must be above the first row of data.

Step 2 Select **Stat>Basic Statistics> 1-sample z.** Press the F3 key to reset the information in the dialog box left over from the previous example.

Step 3 Select **One or more samples, each in a column,** if not already selected. To select the data, click inside the next dialog box; then select C1 SAT Test from the list.

Step 4 Type in the population standard deviation of **100.**

Step 5 Select the box for Perform hypothesis test, then type in the Hypothesized value of **515.**

Step 6 Click the button for [**Options**].

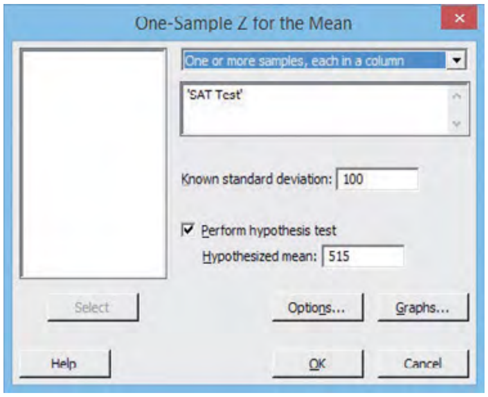
a) Type the default confidence level that is $1 - \alpha$, **0.90** or **90.**

b) Click the drop-down menu for the Alternative hypothesis, 'less than'.

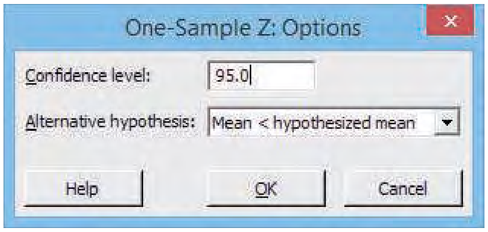
c) Click [**OK**].

Optional: Since there are data, select the [**Graphs**] button; then choose one or more of the three graphs such as the boxplot.

Step 7 Click [**OK**] twice.



One or more samples, each in a column
'SAT Test'
Known standard deviation: 100
☑ Perform hypothesis test
Hypothesized mean: 515



Confidence level: 95.0
Alternative hypothesis: Mean < hypothesized mean

```
One-Sample Z: SAT Test

Test of μ = 515 vs < 515
The assumed standard deviation = 100

                                                90% Upper
Variable    N    Mean   StdDev   SE Mean            Bound       Z       P
SAT Test   36   509.0     92.7      16.7            530.4   −0.36   0.360
```

The test statistic *z* is equal to −0.36. The null hypothesis cannot be rejected since the *P*-value of 0.360 is larger than 0.10.

8–3 *t* Test for a Mean

OBJECTIVE 6

Test means when σ is unknown, using the t test.

When the population standard deviation is unknown, the z test is not typically used for testing hypotheses involving means. A different test, called the *t test,* is used. The distribution of the variable should be approximately normal.

As stated in Chapter 7, the t distribution is similar to the standard normal distribution in the following ways.

1. It is bell-shaped.
2. It is symmetric about the mean.
3. The mean, median, and mode are equal to 0 and are located at the center of the distribution.
4. The curve approaches but never touches the x axis.

The t distribution differs from the standard normal distribution in the following ways.

1. The variance is greater than 1.
2. The t distribution is a family of curves based on the *degrees of freedom,* which is a number related to sample size. (Recall that the symbol for degrees of freedom is d.f. See Section 7–2 for an explanation of degrees of freedom.)
3. As the sample size increases, the t distribution approaches the normal distribution.

The t test is defined next.

The ***t* test** is a statistical test for the mean of a population and is used when the population is normally or approximately normally distributed and σ is unknown.

The formula for the t test is

$$t = \frac{\overline{X} - \mu}{s/\sqrt{n}}$$

The degrees of freedom are d.f. $= n - 1$.

The formula for the t test is similar to the formula for the z test. But since the population standard deviation σ is unknown, the sample standard deviation s is used instead.

The critical values for the t test are given in Table F in Appendix A. For a one-tailed test, find the α level by looking at the top row of the table and finding the appropriate column. Find the degrees of freedom by looking down the left-hand column.

Notice that the degrees of freedom are given for values from 1 through 30, then at intervals above 30. When the degrees of freedom are above 30, some texts will tell you to use the nearest table value; however, in this text, you should always round down to the nearest table value. For example, if d.f. = 59, use d.f. = 55 to find the critical value or values. This is a conservative approach.

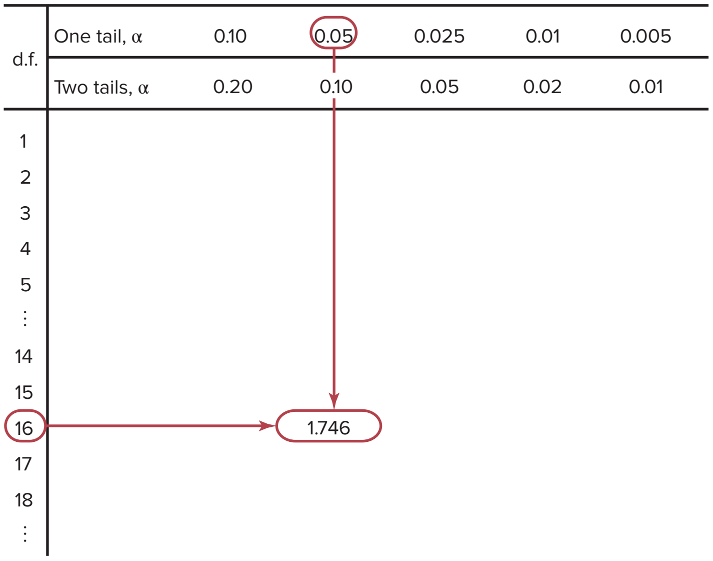
As the degrees of freedom get larger, the critical values approach the z values. Hence, the bottom values (large sample size) are the same as the z values that were used in the last section.

EXAMPLE 8–8

Find the critical t value for $\alpha = 0.05$ with d.f. = 16 for a right-tailed t test.

SOLUTION

Find the 0.05 column in the top row labeled One tail and 16 in the left-hand column. Where the row and column meet, the appropriate critical value is found; it is +1.746. See Figure 8–19.



d.f.	One tail, α	0.10	0.05	0.025	0.01	0.005
	Two tails, α	0.20	0.10	0.05	0.02	0.01
1						
2						
3						
4						
5						
⋮						
14						
15						
16			1.746			
17						
18						
⋮						

FIGURE 8–19
Finding the Critical Value for the *t* Test in Table F (Example 8–8)

EXAMPLE 8–9

Find the critical *t* value for $\alpha = 0.01$ with d.f. = 24 for a left-tailed *t* test.

SOLUTION

Find the critical value in the 0.01 column in the row labeled One tail, and find 24 in the left column (d.f.). The critical value is -2.492 since the test is left-tailed.

EXAMPLE 8–10

Find the critical values for $\alpha = 0.10$ with d.f. = 18 for a two-tailed *t* test.

SOLUTION

Find the 0.10 column in the row labeled Two tails, and find 18 in the column labeled d.f. The critical values are $+1.734$ and -1.734.

EXAMPLE 8–11

Find the critical value for $\alpha = 0.05$ with d.f. = 12 for a right-tailed *t* test.

SOLUTION

Find the critical value in the 0.05 column in the One-tail row, and 12 in the d.f. column. The critical value is $+1.782$.

In order to test hypotheses regarding the population mean when the standard deviation is unknown, these assumptions should be verified first.

Assumptions for the *t* Test for a Mean When σ Is Unknown

1. The sample is a random sample.
2. Either $n \geq 30$ or the population is normally distributed when $n < 30$.

In this text, the assumptions will be stated in the exercises; however, when encountering statistics in other situations, you must check to see that these assumptions have been met before proceeding.

When you test hypotheses by using the t test (traditional method), follow the same procedure as for the z test, except use Table F.

Step 1 State the hypotheses and identify the claim.

Step 2 Find the critical value(s) from Table F.

Step 3 Compute the test value.

Step 4 Make the decision to reject or not reject the null hypothesis.

Step 5 Summarize the results.

Remember that the t test should be used when the population is approximately normally distributed and the population standard deviation is unknown.

Examples 8–12 and 8–13 illustrate the application of the t test, using the traditional approach.

EXAMPLE 8–12 Hospital Infections

A medical investigation claims that the average number of infections per week at a hospital in southwestern Pennsylvania is 16.3. A random sample of 10 weeks had a mean number of 17.7 infections. The sample standard deviation is 1.8. Is there enough evidence to reject the investigator's claim at $\alpha = 0.05$? Assume the variable is normally distributed.

Source: Based on information obtained from Pennsylvania Health Care Cost Containment Council.

SOLUTION

Step 1 H_0: $\mu = 16.3$ (claim) and H_1: $\mu \neq 16.3$.

Step 2 The critical values are +2.262 and −2.262 for $\alpha = 0.05$ and d.f. = 9.

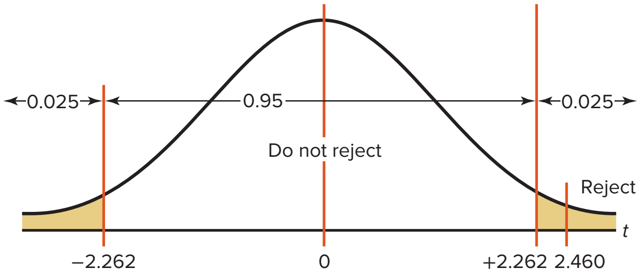
Step 3 The test value is

$$t = \frac{\overline{X} - \mu}{s/\sqrt{n}} = \frac{17.7 - 16.3}{1.8/\sqrt{10}} = 2.460$$

Step 4 Reject the null hypothesis since 2.460 > 2.262. See Figure 8–20.

FIGURE 8–20

Summary of the t Test of Example 8–12



Step 5 There is enough evidence to reject the claim that the average number of infections is 16.3.

EXAMPLE 8–13 Aspirin Consumption

Medical researchers found that people who take aspirin take on average 8.6 tablets a week. A medical doctor wanted to see if he could reduce the number of aspirin tablets a person consumes if the subjects practiced yoga twice a week. A random sample of 12 people who take aspirin were taught yoga, and after a 12-week period the doctor found that they took the following number of tablets for one week:

7 4 6 8 9 9 5 6 2 7 10 9

Test the claim at $\alpha = 0.10$. Assume the variable is normally distributed.

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon \mu = 8.6 \qquad \text{and} \qquad H_1\colon \mu < 8.6 \text{ (claim)}$$

Step 2 Find the critical value. At $\alpha = 0.10$ and d.f. = 11, the critical value is -1.363.

Step 3 Compute the test value. In this case you must find the mean and standard deviation for the data. Use the formulas in Chapter 3 or your calculator.

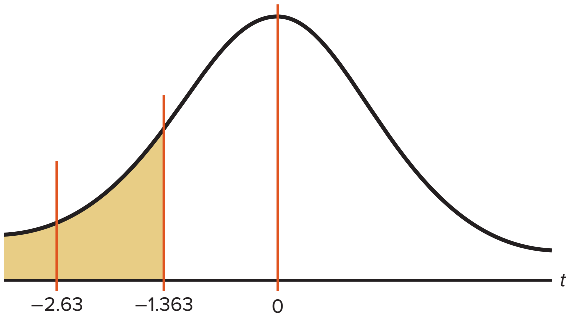
$$\overline{X} = 6.8 \qquad s = 2.37$$

$$t = \frac{\overline{X} - \mu}{s/\sqrt{n}} = \frac{6.8 - 8.6}{2.37/\sqrt{12}} = -2.63$$

Step 4 Make the decision. Reject the null hypothesis since -2.63 falls in the critical region. See Figure 8–21.

FIGURE 8–21

Critical Value and Test Value for Example 8–13



Step 5 Summarize the results. There is enough evidence to support the claim that practicing yoga reduces the weekly aspirin intake.

The *P*-values for the *t* test can be found by using Table F; however, specific *P*-values for *t* tests cannot be obtained from the table since only selected values of α (for example, 0.01, 0.05) are given. To find specific *P*-values for *t* tests, you would need a table similar to Table E for each degree of freedom. Since this is not practical, only *intervals* can be found for *P*-values. Examples 8–14 to 8–16 show how to use Table F to determine intervals for *P*-values for the *t* test.

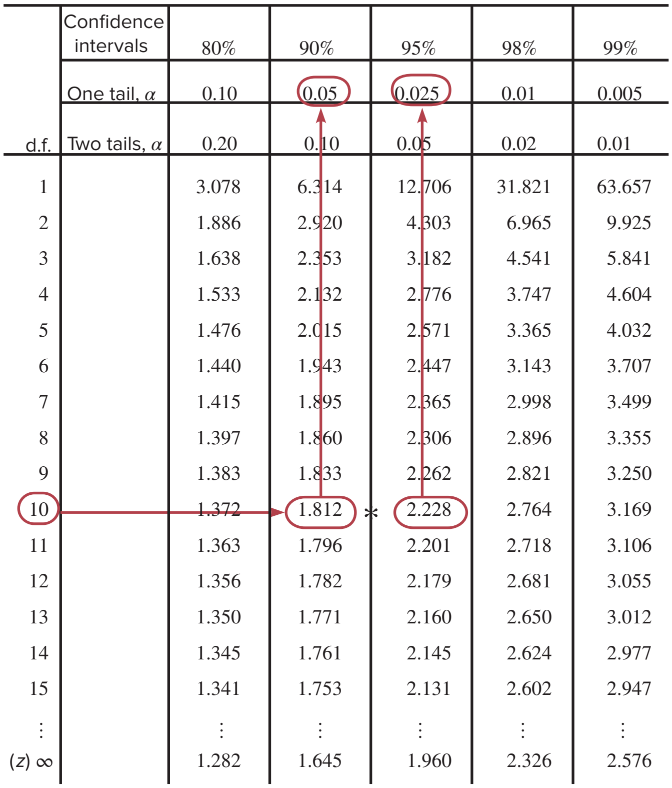
EXAMPLE 8–14

Find the *P*-value when the *t* test value is 2.056, the sample size is 11, and the test is right-tailed.

SOLUTION

To get the *P*-value, look across the row with 10 degrees of freedom (d.f. $= n - 1$) in Table F and find the two values that 2.056 falls between. They are 1.812 and 2.228. Since this is a right-tailed test, look up to the row labeled One tail, α and find the two α values corresponding to 1.812 and 2.228. They are 0.05 and 0.025, respectively. See Figure 8–22.

FIGURE 8–22
Finding the *P*-Value for Example 8–14



	Confidence intervals	80%	90%	95%	98%	99%
	One tail, α	0.10	0.05	0.025	0.01	0.005
d.f.	Two tails, α	0.20	0.10	0.05	0.02	0.01
1		3.078	6.314	12.706	31.821	63.657
2		1.886	2.920	4.303	6.965	9.925
3		1.638	2.353	3.182	4.541	5.841
4		1.533	2.132	2.776	3.747	4.604
5		1.476	2.015	2.571	3.365	4.032
6		1.440	1.943	2.447	3.143	3.707
7		1.415	1.895	2.365	2.998	3.499
8		1.397	1.860	2.306	2.896	3.355
9		1.383	1.833	2.262	2.821	3.250
10		1.372	1.812 *	2.228	2.764	3.169
11		1.363	1.796	2.201	2.718	3.106
12		1.356	1.782	2.179	2.681	3.055
13		1.350	1.771	2.160	2.650	3.012
14		1.345	1.761	2.145	2.624	2.977
15		1.341	1.753	2.131	2.602	2.947
⋮		⋮	⋮	⋮	⋮	⋮
(z) ∞		1.282	1.645	1.960	2.326	2.576

*2.056 falls between 1.812 and 2.228.

Hence, the *P*-value would be contained in the interval $0.025 < P\text{-value} < 0.05$. This means that the *P*-value is between 0.025 and 0.05. If α were 0.05, you would reject the null hypothesis since the *P*-value is less than 0.05. But if α were 0.01, you would not reject the null hypothesis since the *P*-value is greater than 0.01. (Actually, it is greater than 0.025.)

EXAMPLE 8–15

Find the *P*-value when the *t* test value is 2.983, the sample size is 6, and the test is two-tailed.

SOLUTION

To get the *P*-value, look across the row with d.f. = 5 in Table F and find the two values that 2.983 falls between. They are 2.571 and 3.365. Then look up the row labeled Two tails, α to find the corresponding α values.

In this case, they are 0.05 and 0.02. Hence, the *P*-value is contained in the interval $0.02 < P\text{-value} < 0.05$. This means that the *P*-value is between 0.02 and 0.05. In this case, if $\alpha = 0.05$, the null hypothesis can be rejected since $P\text{-value} < 0.05$; but if $\alpha = 0.01$, the null hypothesis cannot be rejected since $P\text{-value} > 0.01$ (actually $P\text{-value} > 0.02$).

***Note:* Since many of you will be using calculators or computer programs that give the specific *P*-value for the *t* test and other tests presented later in this textbook, these specific values, in addition to the intervals, will be given for the answers to the examples and exercises.**

The *P*-value obtained from a calculator for Example 8–14 is 0.033. The *P*-value obtained from a calculator for Example 8–15 is 0.031.

To test hypotheses using the *P*-value method, follow the same steps as explained in Section 8–2. These steps are repeated here.

Step 1 State the hypotheses and identify the claim.
Step 2 Compute the test value.
Step 3 Find the *P*-value.
Step 4 Make the decision.
Step 5 Summarize the results.

This method is shown in Example 8–16.

EXAMPLE 8–16 Jogger's Oxygen Uptake

A physician claims that joggers' maximal volume oxygen uptake is greater than the average of all adults. A random sample of 15 joggers has a mean of 40.6 milliliters per kilogram (ml/kg) and a standard deviation of 6 ml/kg. If the average of all adults is 36.7 ml/kg, is there enough evidence to support the physician's claim at $\alpha = 0.05$? Assume the variable is normally distributed.

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon \mu = 36.7 \qquad \text{and} \qquad H_1\colon \mu > 36.7 \text{ (claim)}$$

Step 2 Compute the test value. The test value is

$$t = \frac{\overline{X} - \mu}{s/\sqrt{n}} = \frac{40.6 - 36.7}{6/\sqrt{15}} = 2.517$$

Step 3 Find the *P*-value. Looking across the row with d.f. = 14 in Table F, you see that 2.517 falls between 2.145 and 2.624, corresponding to $\alpha = 0.025$ and $\alpha = 0.01$ since this is a right-tailed test. Hence, *P*-value > 0.01 and *P*-value < 0.025, or 0.01 < *P*-value < 0.025. That is, the *P*-value is somewhere between 0.01 and 0.025. (The *P*-value obtained from a calculator is 0.012.)

Step 4 Reject the null hypothesis since *P*-value < 0.05 (that is, *P*-value < α).

Step 5 There is enough evidence to support the claim that the joggers' maximal volume oxygen uptake is greater than 36.7 ml/kg.

Interesting Fact

The area of Alaska contains $\frac{1}{6}$ of the total area of the United States.

Students sometimes have difficulty deciding whether to use the *z* test or *t* test. The rules are the same as those pertaining to confidence intervals.

1. If σ is known, use the *z* test. The variable must be normally distributed if $n < 30$.
2. If σ is unknown but $n \geq 30$, use the *t* test.
3. If σ is unknown and $n < 30$, use the *t* test. (The population must be approximately normally distributed.)

These rules are summarized in Figure 8–23.

SPEAKING OF STATISTICS Can Sunshine Relieve Pain?

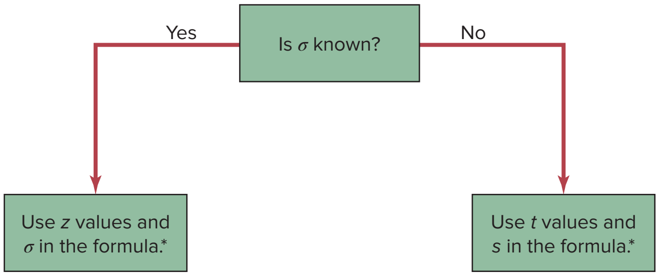
A study conducted at the University of Pittsburgh showed that hospital patients in rooms with lots of sunlight required less pain medication the day after surgery and during their total stay in the hospital than patients who were in darker rooms.

Patients in the sunny rooms averaged 3.2 milligrams of pain reliever per hour for their total stay as opposed to 4.1 milligrams per hour for those in darker rooms. This study compared two groups of patients. Although no statistical tests were mentioned in the article, what statistical test do you think the researchers used to compare the groups?

© Image Source/Getty Images RF

FIGURE 8–23

Using the z or t Test


Is σ known?
Yes
No
Use z values and σ in the formula.*
Use t values and s in the formula.*

*If $n < 30$, the variable must be normally distributed.

Applying the Concepts 8–3

How Much Nicotine Is in Those Cigarettes?

A tobacco company claims that its best-selling cigarettes contain at most 40 mg of nicotine. This claim is tested at the 1% significance level by using the results of 15 randomly selected cigarettes. The mean is 42.6 mg and the standard deviation is 3.7 mg. Evidence suggests that nicotine is normally distributed. Information from a computer output of the hypothesis test is listed.

Sample mean = 42.6	P-value = 0.008
Sample standard deviation = 3.7	Significance level = 0.01
Sample size = 15	Test statistic $t = 2.72155$
Degrees of freedom = 14	Critical value $t = 2.62449$

1. What are the degrees of freedom?
2. Is this a z or t test?
3. Is this a comparison of one or two samples?
4. Is this a right-tailed, left-tailed, or two-tailed test?
5. From observing the P-value, what would you conclude?
6. By comparing the test statistic to the critical value, what would you conclude?
7. Is there a conflict in this output? Explain.
8. What has been proved in this study?

See page 486 for the answers.

Exercises 8–3

1. In what ways is the t distribution similar to the standard normal distribution? In what ways is the t distribution different from the standard normal distribution?

2. What are the degrees of freedom for the t test?

3. Find the critical value (or values) for the t test for each.

a. $n = 12$, $\alpha = 0.01$, left-tailed
b. $n = 16$, $\alpha = 0.05$, right-tailed
c. $n = 7$, $\alpha = 0.10$, two-tailed
d. $n = 11$, $\alpha = 0.025$, right-tailed
e. $n = 10$, $\alpha = 0.05$, two-tailed

4. Find the critical value (or values) for the t test for each.

a. $n = 15$, $\alpha = 0.05$, right-tailed
b. $n = 23$, $\alpha = 0.005$, left-tailed
c. $n = 28$, $\alpha = 0.01$, two-tailed
d. $n = 17$, $\alpha = 0.02$, two-tailed

5. Using Table F, find the P-value interval for each test value.

a. $t = 2.321$, $n = 15$, right-tailed
b. $t = 1.945$, $n = 28$, two-tailed
c. $t = -1.267$, $n = 8$, left-tailed
d. $t = 1.562$, $n = 17$, two-tailed

6. Using Table F, find the P-value interval for each test value.

a. $t = 3.025$, $n = 24$, right-tailed
b. $t = -1.145$, $n = 5$, left-tailed
c. $t = 2.179$, $n = 13$, two-tailed
d. $t = 0.665$, $n = 10$, right-tailed

For Exercises 7 through 23, perform each of the following steps.

a. State the hypotheses and identify the claim.
b. Find the critical value(s).
c. Find the test value.
d. Make the decision.
e. Summarize the results.

Use the traditional method of hypothesis testing unless otherwise specified.

Assume that the population is approximately normally distributed.

7. Cigarette Smoking A researcher found that a cigarette smoker smokes on average 31 cigarettes a day. She feels that this average is too high. She selected a random sample of 10 smokers and found that the mean number of cigarettes they smoked per day was 28. The sample standard deviation was 2.7. At $\alpha = 0.05$ is there enough evidence to support her claim?

8. Cost of Braces The average cost for teeth straightening with metal braces is approximately \$5400. A nationwide franchise thinks that its cost is below that figure. A random sample of 28 patients across the country had an average cost of \$5250 with a standard deviation of \$629. At $\alpha = 0.025$, can it be concluded that the mean is less than \$5400?

9. Heights of Tall Buildings A researcher estimates that the average height of the buildings of 30 or more stories in a large city is at least 700 feet. A random sample of 10 buildings is selected, and the heights in feet are shown. At $\alpha = 0.025$, is there enough evidence to reject the claim?

485 511 841 725 615
520 535 635 616 582

Source: Pittsburgh Tribune-Review.

10. Number of Words in a Novel The National Novel Writing Association states that the average novel is at least 50,000 words. A particularly ambitious writing club at a college-preparatory high school had randomly selected members with works of the following lengths. At $\alpha = 0.10$, is there sufficient evidence to conclude that the mean length is greater than 50,000 words?

48,972 50,100 51,560
49,800 50,020 49,900
52,193

11. Television Viewing by Teens Teens are reported to watch the fewest total hours of television per week of all the demographic groups. The average television viewing for teens on Sunday from 1:00 to 7:00 P.M. is 58 minutes. A random sample of local teens disclosed the following times for Sunday afternoon television viewing. At $\alpha = 0.01$, can it be concluded that the average is greater than the national viewing time? (*Note:* Change all times to minutes.)

2:30 2:00 1:30 3:20
1:00 2:15 1:50 2:10
1:30 2:30

Source: World Almanac.

12. Chocolate Chip Cookie Calories The average 1-ounce chocolate chip cookie contains 110 calories. A random sample of 15 different brands of 1-ounce chocolate chip cookies resulted in the following calorie amounts. At the $\alpha = 0.01$ level, is there sufficient evidence that the average calorie content is greater than 110 calories?

100 125 150 160 185 125 155 145 160
100 150 140 135 120 110

Source: The Doctor's Pocket Calorie, Fat, and Carbohydrate Counter.

13. Sleep Time A person read that the average number of hours an adult sleeps on Friday night to Saturday morning was 7.2 hours. The researcher feels that college students do not sleep 7.2 hours on average. The researcher randomly selected 15 students and found that on average they slept 8.3 hours. The standard deviation of the sample is 1.2 hours. At $\alpha = 0.05$, is there enough evidence to say that college students do not sleep 7.2 hours on average?

14. **Internet Visits** A U.S. Web Usage Snapshot indicated a monthly average of 36 Internet visits a particular website per user from home. A random sample of 24 Internet users yielded a sample mean of 42.1 visits with a standard deviation of 5.3. At the 0.01 level of significance, can it be concluded that this differs from the national average?

    Source: New York Times Almanac.

15. **Cell Phone Bills** The average monthly cell phone bill was reported to be \$50.07 by the U.S. Wireless Industry. Random sampling of a large cell phone company found the following monthly cell phone charges (in dollars):

    55.83 49.88 62.98 70.42
    60.47 52.45 49.20 50.02
    58.60 51.29

    At the 0.05 level of significance, can it be concluded that the average phone bill has increased?

    Source: World Almanac.

16. **Teaching Assistants' Stipends** A random sample of stipends of teaching assistants in economics is listed. Is there sufficient evidence at the $\alpha = 0.05$ level to conclude that the average stipend differs from \$15,000? The stipends listed (in dollars) are for the academic year.

    14,000 18,000 12,000 14,356 13,185
    13,419 14,000 11,981 17,604 12,283
    16,338 15,000

    Source: Chronicle of Higher Education.

17. **Medical Operations** The director of a medical hospital feels that her surgeons perform fewer operations per year than the national average of 211. She selected a random sample of 15 surgeons and found that the mean number of operations they performed was 208.8. The standard deviation of the sample was 3.8. Is there enough evidence to support the director's feelings at $\alpha = 0.10$? Would the null hypothesis be rejected at $\alpha = 0.01$?

18. **Cell Phone Call Lengths** The average local cell phone call length was reported to be 2.27 minutes. A random sample of 20 phone calls showed an average of 2.98 minutes in length with a standard deviation of 0.98 minute. At $\alpha = 0.05$, can it be concluded that the average differs from the population average?

    Source: World Almanac.

19. **Commute Time to Work** A survey of 15 large U.S. cities finds that the average commute time one way is 25.4 minutes. A chamber of commerce executive feels that the commute in his city is less and wants to publicize this. He randomly selects 25 commuters and finds the average is 22.1 minutes with a standard deviation of 5.3 minutes. At $\alpha = 0.10$, is he correct?

    Source: New York Times Almanac.

20. **Average Family Size** The average family size was reported as 3.18. A random sample of families in a particular school district resulted in the following family sizes:

    5 4 5 4 4 3 6 4 3 3 5
    6 3 3 2 7 4 5 2 2 2 3
    5 2

    At $\alpha = 0.05$, does the average family size differ from the national average?

    Source: New York Times Almanac.

21. **Doctor Visits** A report by the Gallup Poll stated that on average a woman visits her physician 5.8 times a year. A researcher randomly selects 20 women and obtained these data.

    3 2 1 3 7 2 9 4 6 6
    8 0 5 6 4 2 1 3 4 1

    At $\alpha = 0.05$, can it be concluded that the average is still 5.8 visits per year? Use the *P*-value method.

22. **Number of Jobs** The U.S. Bureau of Labor and Statistics reported that a person between the ages of 18 and 34 has had an average of 9.2 jobs. To see if this average is correct, a researcher selected a random sample of 8 workers between the ages of 18 and 34 and asked how many different places they had worked. The results were as follows:

    8 12 15 6 1 9 13 2

    At $\alpha = 0.05$, can it be concluded that the mean is 9.2? Use the *P*-value method. Give one reason why the respondents might not have given the exact number of jobs that they have worked.

23. **Water Consumption** The *Old Farmer's Almanac* stated that the average consumption of water per person per day was 123 gallons. To test the hypothesis that this figure may no longer be true, a researcher randomly selected 16 people and found that they used on average 119 gallons per day and $s = 5.3$. At $\alpha = 0.05$, is there enough evidence to say that the *Old Farmer's Almanac* figure might no longer be correct? Use the *P*-value method.

Technology Step by Step

TI-84 Plus
Step by Step

Hypothesis Test for the Mean and the *t* Distribution (Data)

1. Enter the data values into L_1.
2. Press **STAT** and move the cursor to TESTS.
3. Press **2** for T-Test.
4. Move the cursor to Data and press **ENTER.**
5. Type in the appropriate values.

6. Move the cursor to the appropriate alternative hypothesis and press **ENTER.**
7. Move the cursor to Calculate and press **ENTER.**

Hypothesis Test for the Mean and the *t* Distribution (Statistics)

1. Press **STAT** and move the cursor to TESTS.
2. Press **2** for T-Test.
3. Move the cursor to Stats and press **ENTER.**
4. Type in the appropriate values.
5. Move the cursor to the appropriate alternative hypothesis and press **ENTER.**
6. Move the cursor to Calculate and press **ENTER.**

EXCEL
Step by Step

Hypothesis Test for the Mean: *t* Test

Excel does not have a procedure to conduct a hypothesis test for the mean. However, you may conduct the test of the mean using the MegaStat Add-in available in your online resources. If you have not installed this add-in, do so, following the instructions from the Chapter 1 Excel Step by Step.

Substitute Teachers' Salaries

An educator claims that the average salary of substitute teachers in school districts in Allegheny County, Pennsylvania, is less than $60 per day. A random sample of eight school districts is selected, and the daily salaries (in dollars) are shown. Is there enough evidence to support the educator's claim at $\alpha = 0.10$? Assume the variable is normally distributed.

60 56 60 55 70 55 60 55

Source: Pittsburgh Tribune-Review.

The MegaStat *t* test uses the *P*-value method. Therefore, it is not necessary to enter a significance level.

1. Enter the data into column A of a new worksheet.
2. From the toolbar, select Add-Ins, **MegaStat>Hypothesis Tests>Mean vs. Hypothesized Value.** *Note:* You may need to open MegaStat from the MegaStat.xls file on your computer's hard drive.
3. Select data input and type **A1:A8** as the Input Range.
4. Type **60** for the Hypothesized mean and select the "less than" Alternative.
5. Select *t* test and click [OK].

The result of the procedure is shown next.

Hypothesis Test: Mean vs. Hypothesized Value

```
60.000   Hypothesized value
58.875   Mean data
 5.083   Standard deviation
 1.797   Standard error
     8   n
     7   d.f.

 –0.63   t
0.2756   P-value (one-tailed, lower)
```

MINITAB
Step by Step

Hypothesis Test for the Mean and the *t* Distribution

MINITAB can be used to look up a critical value of *t*.

Example 8–8

Find the critical *t* value for $\alpha = 0.05$ with d.f. = 16 for a right-tailed *t* test.

Step 1 To find the critical value of *t* for a right-tailed test, select **Graph>Probability Distribution Plot,** then **View Probability,** and then click [**OK**].

Step 2 Change the Distribution to a *t* distribution with Degrees of freedom equal to **16.**

Step 3 Click the tab for **Shaded Area.**

- a) Select the ratio button for **Probability.**
- b) Select **Right Tail.**
- c) Type in the value of alpha for Probability, **0.05.**
- d) Click **[OK].**

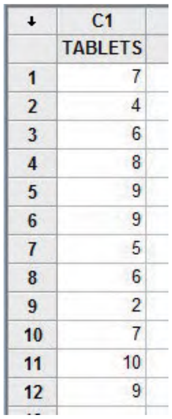
The critical value of *t* to three decimal places is 1.746.

You may click the Edit Last Dialog button and then change the settings for additional critical values.

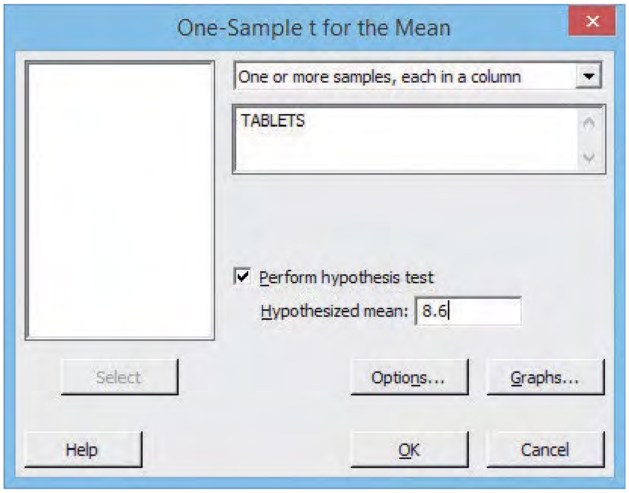
Example 8–13 Aspirin Tablets

MINITAB will calculate the test statistic and *P*-value from the data.

Step 1 Type the data into a new MINITAB worksheet. All 12 values must be in C1. The label must be above the first row of data. Do not type the commas in large numbers.



↓	C1
	TABLETS
1	7
2	4
3	6
4	8
5	9
6	9
7	5
8	6
9	2
10	7
11	10
12	9



One-Sample t for the Mean

One or more samples, each in a column

TABLETS

☑ Perform hypothesis test

Hypothesized mean: 8.6

Step 2 Select **Stat>Basic Statistics> 1-sample t.**

Step 3 Select **One or more samples, each in a column,** if not already selected.

To select the data, click inside the next dialog box; then select C1 TABLETS from the list.

Step 4 Select the box for Perform hypothesis test, and then type in the Hypothesized value of **8.6.**

Step 5 Click the button for **[Options].**

- a) Type the default confidence level that is $1 - \alpha = 0.90$ or 90.
- b) Click the drop down menu for the Alternative hypothesis, **"Mean < hypothesized mean."**
- c) Click **[OK].**

  Optional: Since there are data, select the **[Graphs]** button and then choose one or more of the three graphs such as the boxplot.

Step 6 Click **[OK]** twice.

```
One-Sample t: Aspirin Tablets
Test of μ = 8.6 vs < 8.6
                                                  90% Upper
Variable        N     Mean    StDev    SE Mean        Bound       T       P
TABLETS        12    6.833    2.368      0.683        7.765   –2.58   0.013
```

The test statistic is –2.58. Since the *P*-value of 0.013 is less than alpha, the null hypothesis is rejected.

8–4 *z* Test for a Proportion

OBJECTIVE ⑦

Test proportions, using the *z* test.

Many hypothesis-testing situations involve proportions. Recall from Chapter 7 that a *proportion* is the same as a percentage of the population.

These data were obtained from *The Book of Odds* by Michael D. Shook and Robert L. Shook (New York: Penguin Putnam, Inc.):

- 59% of consumers purchase gifts for their fathers.
- 85% of people over 21 said they have entered a sweepstakes.
- 51% of Americans buy generic products.
- 35% of Americans go out for dinner once a week.

A hypothesis test involving a population proportion can be considered as a binomial experiment when there are only two outcomes and the probability of a success does not change from trial to trial. Recall from Section 5–3 that the mean is $\mu = np$ and the standard deviation is $\sigma = \sqrt{npq}$ for the binomial distribution.

Since a normal distribution can be used to approximate the binomial distribution when $np \geq 5$ and $nq \geq 5$, the standard normal distribution can be used to test hypotheses for proportions.

Formula for the *z* Test for Proportions

$$z = \frac{\hat{p} - p}{\sqrt{pq/n}}$$

where $\hat{p} = \frac{X}{n}$ sample proportion
$p =$ population proportion
$n =$ sample size

The formula is derived from the normal approximation to the binomial and follows the general formula

$$\text{Test value} = \frac{(\text{observed value}) - (\text{expected value})}{\text{standard error}}$$

We obtain $\hat{p}$ from the sample (i.e., observed value), p is the expected value (i.e., hypothesized population proportion), and $\sqrt{pq/n}$ is the standard error.

The formula $z = \dfrac{\hat{p} - p}{\sqrt{pq/n}}$ can be derived from the formula $z = \dfrac{X - \mu}{\sigma}$ by substituting $\mu = np$ and $\sigma = \sqrt{npq}$ and then dividing both numerator and denominator by n. Some algebra is used. See Exercise 23 in this section.

The assumptions for testing a proportion are given next.

Assumptions for Testing a Proportion

1. The sample is a random sample.
2. The conditions for a binomial experiment are satisfied. (See Chapter 5.)
3. $np \geq 5$ and $nq \geq 5$.

In this book, the assumptions will be stated in the exercises; however, when encountering statistics in other situations, you must check to see that these assumptions have been met before proceeding.

The steps for hypothesis testing are the same as those shown in Section 8–3. Table E is used to find critical values and *P*-values.

Examples 8–17 to 8–19 show the traditional method of hypothesis testing. Example 8–20 shows the *P*-value method.

Sometimes it is necessary to find $\hat{p}$, as shown in Examples 8–17, 8–19, and 8–20, and sometimes $\hat{p}$ is given in the exercise. See Example 8–18.

EXAMPLE 8–17 Obese Young People

A researcher claims that based on the information obtained from the Centers for Disease Control and Prevention, 17% of young people ages 2–19 are obese. To test this claim, she randomly selected 200 people ages 2–19 and found that 42 were obese. At $\alpha = 0.05$, is there enough evidence to reject the claim?

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon p = 0.17 \text{ (claim)} \qquad \text{and} \qquad H_1\colon p \neq 0.17$$

Step 2 Find the critical values. Since $\alpha = 0.05$ and the test is two-tailed, the critical values are ± 1.96.

Step 3 Compute the test value. First, it is necessary to find $\hat{p}$.

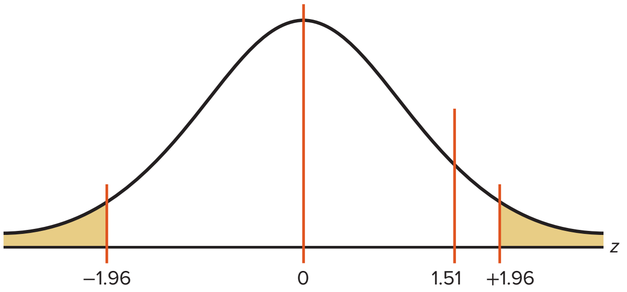
$$\hat{p} = \frac{X}{n} = \frac{42}{200} = 0.21 \qquad p = 0.17 \qquad q = 1 - p = 1 - 0.17 = 0.83$$

Substitute in the formula.

$$z = \frac{\hat{p} - p}{\sqrt{pq/n}} = \frac{0.21 - 0.17}{\sqrt{(0.17)(0.83)/200}} = 1.51$$

Step 4 Make the decision. Do not reject the null hypothesis since the test value falls in the noncritical region. See Figure 8–24.

FIGURE 8–24
Critical and Test Values for Example 8–17



Step 5 Summarize the results. There is not enough evidence to reject the claim that 17% of young people ages 2–19 are obese.

EXAMPLE 8–18 Caesarean Babies

A report stated that 21% of babies are delivered by Caesarean section. A researcher believes that the percentage is less than 21% in the large hospital in his hometown. He randomly selected 50 newborn infants and found that 8 were born by Caesarean section. At $\alpha = 0.01$, is there enough evidence to support his claim?

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon p = 0.21 \qquad \text{and} \qquad H_1\colon p < 0.21 \text{ (claim)}$$

Step 2 Find the critical value. Since $\alpha = 0.01$ and the test is one-tailed, the critical value is -2.33.

Step 3 Compute the test value.

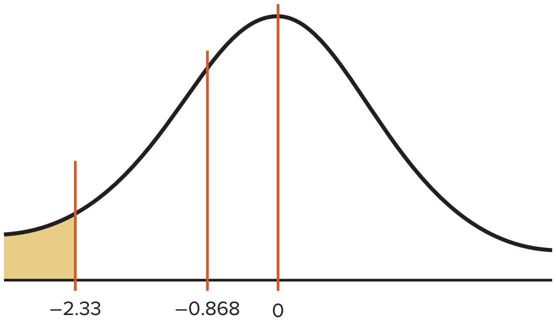
$$\hat{p} = \frac{8}{50} = 0.16 \qquad q = 1 - p = 1 - 0.21 = 0.79$$

Substitute in the formula.

$$z = \frac{\hat{p} - p}{\sqrt{pq/n}} = \frac{0.16 - 0.21}{\sqrt{\dfrac{(0.21)(0.79)}{50}}} = -0.868$$

Step 4 Make the decision. Do not reject the null hypothesis since the critical value does not fall in the critical region. See Figure 8–25.

FIGURE 8–25
Critical and Test Value for Example 8–18



Step 5 Summarize the results. There is not enough evidence to support the claim that the proportion of Caesarian births is less than 21%. The test value does not fall in the rejection region.

EXAMPLE 8–19 Replacing $1 Bills with $1 Coins

A statistician read that at least 77% of the population oppose replacing \$1 bills with \$1 coins. To see if this claim is valid, the statistician selected a random sample of 80 people and found that 55 were opposed to replacing the \$1 bills. At $\alpha = 0.01$, test the claim that at least 77% of the population are opposed to the change.

Source: USA TODAY.

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon p = 0.77 \text{ (claim)} \qquad \text{and} \qquad H_1\colon p < 0.77$$

Step 2 Find the critical value(s). Since $\alpha = 0.01$ and the test is left-tailed, the critical value is −2.33.

Step 3 Compute the test value.

$$\hat{p} = \frac{X}{n} = \frac{55}{80} = 0.6875$$

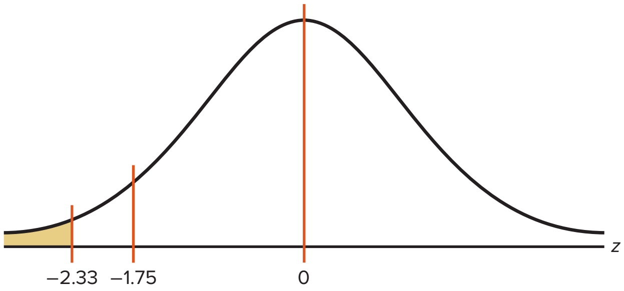
$$p = 0.77 \qquad \text{and} \qquad q = 1 - 0.77 = 0.23$$

$$z = \frac{\hat{p} - p}{\sqrt{pq/n}} = \frac{0.6875 - 0.77}{\sqrt{(0.77)(0.23)/80}} = -1.75$$

Step 4 Do not reject the null hypothesis, since the test value does not fall in the critical region, as shown in Figure 8–26.

FIGURE 8–26

Critical and Test Values for Example 8–19



Step 5 There is not enough evidence to reject the claim that at least 77% of the population oppose replacing $1 bills with $1 coins.

EXAMPLE 8–20 Attorney Advertisements

An attorney claims that more than 25% of all lawyers advertise. A random sample of 200 lawyers in a certain city showed that 63 had used some form of advertising. At $\alpha = 0.05$, is there enough evidence to support the attorney's claim? Use the *P*-value method.

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon p = 0.25 \qquad \text{and} \qquad H_1\colon p > 0.25 \text{ (claim)}$$

Step 2 Compute the test value.

$$\hat{p} = \frac{X}{n} = \frac{63}{200} = 0.315$$

$$p = 0.25 \qquad \text{and} \qquad q = 1 - 0.25 = 0.75$$

$$z = \frac{\hat{p} - p}{\sqrt{pq/n}} = \frac{0.315 - 0.25}{\sqrt{(0.25)(0.75)/200}} = 2.12$$

Step 3 Find the *P*-value. The area under the curve in Table E for $z = 2.12$ is 0.9830. Subtracting the area from 1.0000, you get $1.0000 - 0.9830 = 0.0170$. The *P*-value is 0.0170.

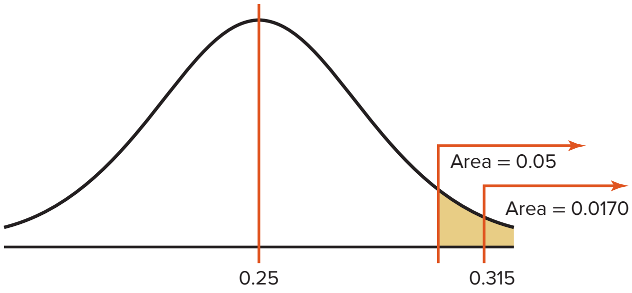
Step 4 Reject the null hypothesis, since $0.0170 < 0.05$ (that is, *P*-value $< \alpha$). See Figure 8–27.

Interesting Facts

Lightning is the second most common killer among storm-related hazards. On average, 73 people are killed each year by lightning. Of people who are struck by lightning, 90% do survive; however, they usually have lasting medical problems or disabilities.

FIGURE 8–27

P-Value and α Value for Example 8–20



Step 5 There is enough evidence to support the attorney's claim that more than 25% of the lawyers use some form of advertising.

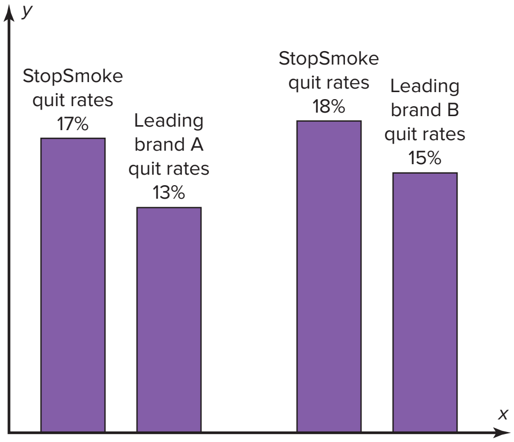
Applying the Concepts 8–4

Quitting Smoking

Assume you are part of a research team that compares products designed to help people quit smoking. Condor Consumer Products Company would like more specific details about the study to be made available to the scientific community. Review the following and then answer the questions about how you would have conducted the study.

New StopSmoke

No method has been proved more effective. StopSmoke provides significant advantages over all other methods. StopSmoke is simpler to use, and it requires no weaning. StopSmoke is also significantly less expensive than the leading brands. StopSmoke's superiority has been proved in two independent studies.



1. What were the statistical hypotheses?
2. What were the null hypotheses?
3. What were the alternative hypotheses?
4. Were any statistical tests run?
5. Were one- or two-tailed tests run?
6. What were the levels of significance?
7. If a type I error was committed, explain what it would have been.
8. If a type II error was committed, explain what it would have been.
9. What did the studies prove?
10. Two statements are made about significance. One states that StopSmoke provides significant advantages, and the other states that StopSmoke is significantly less expensive than other leading brands. Are they referring to statistical significance? What other type of significance is there?

See page 486 for the answers.

Exercises 8–4

1. Give three examples of proportions.

2. Why is a proportion considered a binomial variable?

3. When you are testing hypotheses by using proportions, what are the necessary requirements?

4. What are the mean and the standard deviation of a proportion?

For Exercises 5 through 20, perform each of the following steps.

a. State the hypotheses and identify the claim.
b. Find the critical value(s).
c. Compute the test value.
d. Make the decision.
e. Summarize the results.

Use the traditional method of hypothesis testing unless otherwise specified.

5. Life on Other Planets Forty-six percent of people believe that there is life on other planets in the universe. A scientist does not agree with this finding. He surveyed 120 randomly selected individuals and found 48 believed that there is life on other planets. At $\alpha = 0.10$, is there sufficient evidence to conclude that the percentage differs from 46?

Source: American Health, *Inc.*

6. Stocks and Mutual Fund Ownership It has been found that 50.3% of U.S. households own stocks and mutual funds. A random sample of 300 heads of households indicated that 171 owned some type of stock. At what level of significance would you conclude that this was a significant difference?

Source: www.census.gov

7. **Takeout Food** A magazine article reported that 11% of adults buy takeout food every day. A fast-food restaurant owner surveyed 200 customers and found that 32 said that they purchased takeout food every day. At $\alpha = 0.02$, is there evidence to believe the article's claim? Would the claim be rejected at $\alpha = 0.05$?

8. **Female Physicians** The percentage of physicians who are women is 27.9%. In a survey of physicians employed by a large university health system, 45 of 120 randomly selected physicians were women. Is there sufficient evidence at the 0.05 level of significance to conclude that the proportion of women physicians at the university health system exceeds 27.9%?

   Source: New York Times Almanac.

9. **Runaways** A corrections officer read that 58% of runaways are female. He believes that the percentage is higher than 58. He selected a random sample of 90 runaways and found that 63 were female. At $\alpha = 0.05$, can you conclude that his belief is correct?

   Source: FBI.

10. **Undergraduate Enrollment** It has been found that 85.6% of all enrolled college and university students in the United States are undergraduates. A random sample of 500 enrolled college students in a particular state revealed that 420 of them were undergraduates. Is there sufficient evidence to conclude that the proportion differs from the national percentage? Use $\alpha = 0.05$.

    Source: Time Almanac.

11. **Automobiles Purchased** An automobile owner found that 20 years ago, 76% of Americans said that they would prefer to purchase an American automobile. He believes that the number is much less than 76% today. He selected a random sample of 56 Americans and found that 38 said that they would prefer an American automobile. Can it be concluded that the percentage today is less than 76%? At $\alpha = 0.01$, is he correct?

    Source: Opinion Research Corporation.

12. **Television Set Ownership** According to Nielsen Media Research, of all the U.S. households that owned at least one television set, 83% had two or more sets. A local cable company canvassing the town to promote a new cable service found that of the 300 randomly selected households visited, 240 had two or more television sets. At $\alpha = 0.05$, is there sufficient evidence to conclude that the proportion is less than the one in the report?

    Source: World Almanac.

13. **After-School Snacks** In the *Journal of the American Dietetic Association,* it was reported that 54% of kids said that they had a snack after school. A random sample of 60 kids was selected, and 36 said that they had a snack after school. Use $\alpha = 0.01$ and the P-value method to test the claim. On the basis of the results, should parents be concerned about their children eating a healthy snack?

14. **Natural Gas Heat** The Energy Information Administration reported that 51.7% of homes in the United States were heated by natural gas. A random sample of 200 homes found that 115 were heated by natural gas. Does the evidence support the claim, or has the percentage changed? Use $\alpha = 0.05$ and the P-value method. What could be different if the sample were taken in a different geographic area?

15. **Youth Smoking** Researchers suspect that 18% of all high school students smoke at least one pack of cigarettes a day. At Wilson High School, a randomly selected sample of 300 students found that 50 students smoked at least one pack of cigarettes a day. At $\alpha = 0.05$, test the claim that less than 18% of all high school students smoke at least one pack of cigarettes a day. Use the P-value method.

16. **Exercise to Reduce Stress** A survey by *Men's Health* magazine stated that 14% of men said they used exercise to reduce stress. Use $\alpha = 0.10$. A random sample of 100 men was selected, and 10 said that they used exercise to relieve stress. Use the P-value method to test the claim. Could the results be generalized to all adult Americans?

17. **Borrowing Library Books** For Americans using library services, the American Library Association (ALA) claims that 67% borrow books. A library director feels that this is not true so he randomly selects 100 borrowers and finds that 82 borrowed books. Can he show that the ALA claim is incorrect? Use $\alpha = 0.05$.

    Source: American Library Association; *USA TODAY.*

18. **Doctoral Students' Salaries** Nationally, at least 60% of Ph.D. students have paid assistantships. A college dean feels that this is not true in his state, so he randomly selects 50 Ph.D. students and finds that 26 have assistantships. At $\alpha = 0.05$, is the dean correct?

    Source: U.S. Department of Education, *Chronicle of Higher Education.*

19. **Football Injuries** A report by the NCAA states that 57.6% of football injuries occur during practices. A head trainer claims that this is too high for his conference, so he randomly selects 36 injuries and finds that 17 occurred during practices. Is his claim correct, at $\alpha = 0.05$?

    Source: NCAA Sports Medicine Handbook.

20. **Recycling** Approximately 70% of the U.S. population recycles. According to a green survey of a random sample of 250 college students, 204 said that they recycled. At $\alpha = 0.01$, is there sufficient evidence to conclude that the proportion of college students who recycle is greater than 70%?

Extending the Concepts

When np *or* nq *is not 5 or more, the binomial table (Table B in Appendix A) must be used to find critical values in hypothesis tests involving proportions.*

21. Coin Tossing A coin is tossed 9 times and 3 heads appear. Can you conclude that the coin is not balanced? Use $\alpha = 0.10$. [*Hint:* Use the binomial table and find $2P(X \le 3)$ with $p = 0.5$ and $n = 9$.]

22. First-Class Airline Passengers In the past, 20% of all airline passengers flew first class. In a sample of 15 passengers, 5 flew first class. At $\alpha = 0.10$, can you conclude that the proportions have changed?

23. Show that $z = \dfrac{\hat{p} - p}{\sqrt{pq/n}}$ can be derived from $z = \dfrac{X - \mu}{\sigma}$ by substituting $\mu = np$ and $\sigma = \sqrt{npq}$ and dividing both numerator and denominator by n.

Technology Step by Step

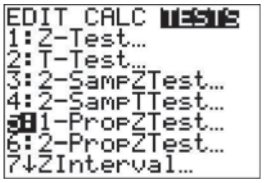
TI-84 Plus
Step by Step

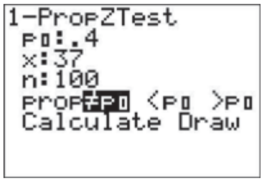
Hypothesis Test for the Proportion

1. Press **STAT** and move the cursor to TESTS.
2. Press **5** for 1-PropZTest.
3. Type in the appropriate values.
4. Move the cursor to the appropriate alternative hypothesis and press **ENTER.**
5. Move the cursor to Calculate and press **ENTER.**

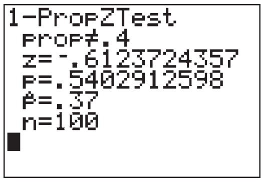
Example

Test the claim that 40% of all telephone customers have call-waiting service, when $n = 100$ and $\hat{p} = 37\%$. Use $\alpha = 0.01$.





```
1-PropZTest
 p0:.4
 x:37
 n:100
 prop≠p0 <p0 >p0
 Calculate Draw
```



```
1-PropZTest
 prop≠.4
 z=-.6123724357
 p=.5402912598
 p̂=.37
 n=100
```

The test statistic is $z = -0.6123724357$, and the P-value is 0.5402912598.

EXCEL
Step by Step

Hypothesis Test for the Proportion: *z* Test

Excel does not have a procedure to conduct a hypothesis test for a single population proportion. However, you may conduct the test of the proportion, using the MegaStat Add-in available in your online resources. If you have not installed this add-in, do so, following the instructions from the Chapter 1 Excel Step by Step.

Example XL8–2

This example relates to the previous example. At the 1% significance level, test the claim that $p = 0.40$. The MegaStat test of the population proportion uses the P-value method. Therefore, it is not necessary to enter a significance level.

1. From the toolbar, select Add-Ins, **MegaStat>Hypothesis Tests>Proportion vs. Hypothesized Value.** *Note:* You may need to open MegaStat from the MegaStat.xls file on your computer's hard drive.
2. Type **0.37** for the Observed proportion, p.
3. Type **0.40** for the Hypothesized proportion, p.
4. Type **100** for the sample size, n.
5. Select the "not equal" Alternative.
6. Click [OK].

The result of the procedure is shown next.

```
Hypothesis Test for Proportion vs. Hypothesized Value
Observed      Hypothesized
    0.37           0.4     p (as decimal)
  37/100        40/100     p (as fraction)
     37.           40.     X
     100           100     n

                 0.049     standard error
                 −0.61     z
                0.5403     p-value (two-tailed)
```

MINITAB
Step by Step

Hypothesis Test for One Proportion and the *z* Distribution

MINITAB can be used to find a critical value of chi-square.

Example 8–17

Test the claim that 17% of young people between the ages of 2 and 19 are obese.
MINITAB will calculate the test statistic and *P*-value based on the normal distribution. There are no data for this example. It doesn't matter what is in the worksheet.

Step 1 Select **Stat>Basic Statistics> 1-proportion.**

Step 2 Select **Summarized data** from the drop down menu.

a) In the dialog box for Number of events, type in the number of successes in the sample, **42.**

b) In the dialog box for Number of trials, type in the sample size, **200.**

Step 3 Select the box for **Perform hypothesis test,** then type the decimal form of the Hypothesized proportion, **0.17.**

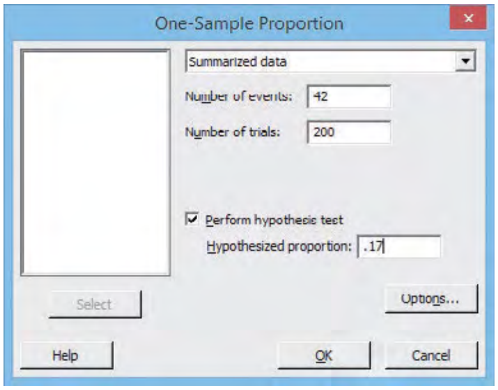
Step 4 Click the button for **[Options].**

a) Type in the confidence level, **95.**

b) The Alternative hypothesis should match the condition in H_1, not equal.

c) Check the box for Use test and interval based on normal distribution.

d) Click [OK] twice.


One-Sample Proportion

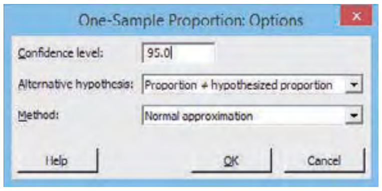
Summarized data

Number of events: 42

Number of trials: 200

☑ Perform hypothesis test

Hypothesized proportion: .17


One-Sample Proportion: Options

Confidence level: 95.0

Alternative hypothesis: Proportion ≠ hypothesized proportion

Method: Normal approximation

In the Session Window the output will include the test statistics, $t = 1.51$ and its *P*-value 0.132. The null hypothesis cannot be rejected.

```
Test and CI for One Proportion
Test of p = 0.17 vs p not = 0.17

Sample    X    N   Sample p         95% CI              Z-Value   P-Value
1        42  200   0.210000   (0.153551, 0.266449)        1.51     0.132

Using the normal approximation.
```

8–5 χ^2 Test for a Variance or Standard Deviation

OBJECTIVE 8

Test variances or standard deviations, using the chi-square test.

In Chapter 7, the chi-square distribution was used to construct a confidence interval for a single variance or standard deviation. This distribution is also used to test a claim about a single variance or standard deviation.

Recall from Chapter 7 the characteristics of the chi-square distribution:

1. All chi-square values are greater than or equal to 0.
2. The chi-square distribution is a family of curves based on the degrees of freedom.
3. The area under each chi-square distribution is equal to 1.
4. The chi-square distributions are positively skewed.

To find the area under the chi-square distribution, use Table G in Appendix A. There are three cases to consider:

1. Finding the chi-square critical value for a specific α when the hypothesis test is right-tailed
2. Finding the chi-square critical value for a specific α when the hypothesis test is left-tailed
3. Finding the chi-square critical values for a specific α when the hypothesis test is two-tailed

Table G is set up so it gives the areas to the right of the critical value; so if the test is right-tailed, just use the area under the α value for the specific degrees of freedom. If the test is left-tailed, subtract the α value from 1; then use the area in the table for that value for a specific d.f. If the test is two-tailed, divide the α value by 2; then use the area under that value for a specific d.f. for the right critical value and the area for the $1 - \alpha/2$ value for the d.f. for the left critical value.

EXAMPLE 8–21

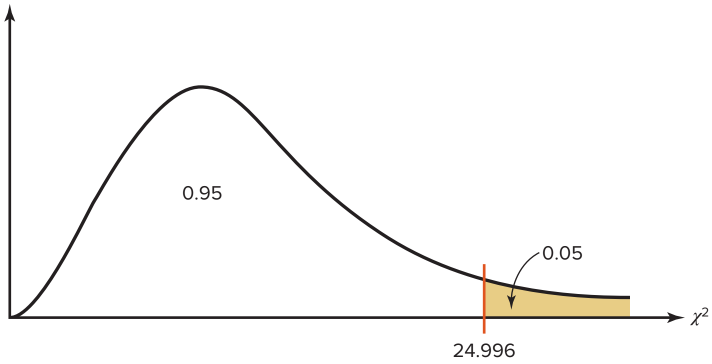
Find the critical chi-square value for 15 degrees of freedom when $\alpha = 0.05$ and the test is right-tailed.

SOLUTION

The distribution is shown in Figure 8–28.

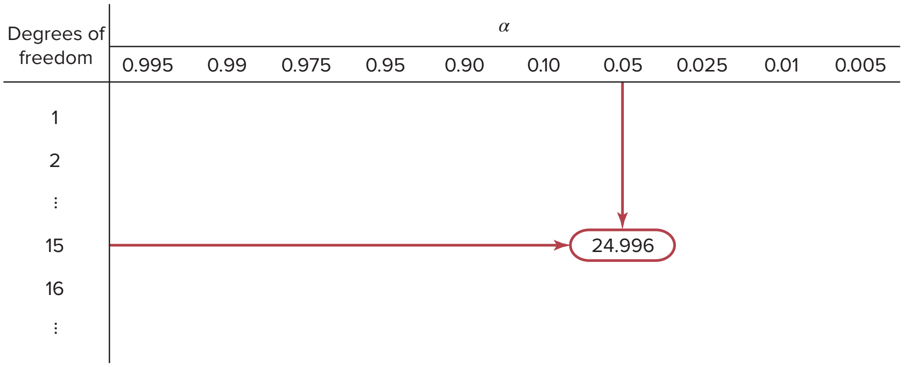
FIGURE 8–28

Chi-Square Distribution for Example 8–21



Find the α value at the top of Table G, and find the corresponding degrees of freedom in the left column. The critical value is located where the two columns meet—in this case, 24.996. See Figure 8–29.

FIGURE 8–29
Locating the Critical Value in Table G for Example 8–21



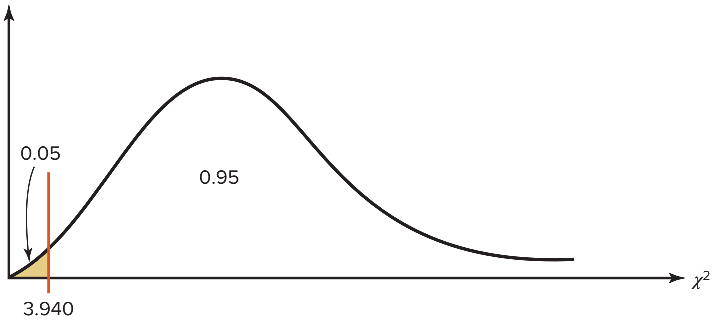
EXAMPLE 8–22

Find the critical chi-square value for 10 degrees of freedom when $\alpha = 0.05$ and the test is left-tailed.

SOLUTION

This distribution is shown in Figure 8–30.

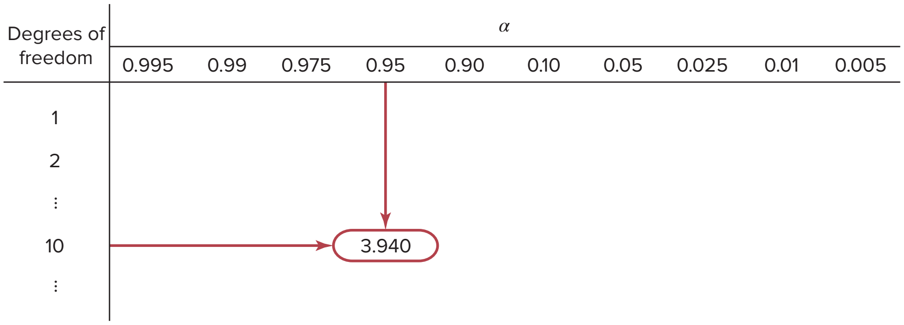
FIGURE 8–30
Chi-Square Distribution for Example 8–22



When the test is left-tailed, the α value must be subtracted from 1, that is, $1 - 0.05 = 0.95$. The left side of the table is used, because the chi-square table gives the area to the right of the critical value, and the chi-square statistic cannot be negative. The table is set up so that it gives the values for the area to the right of the critical value. In this case, 95% of the area will be to the right of the value.

For 0.95 and 10 degrees of freedom, the critical value is 3.940. See Figure 8–31.

FIGURE 8–31
Locating the Critical Value in Table G for Example 8–22



EXAMPLE 8–23

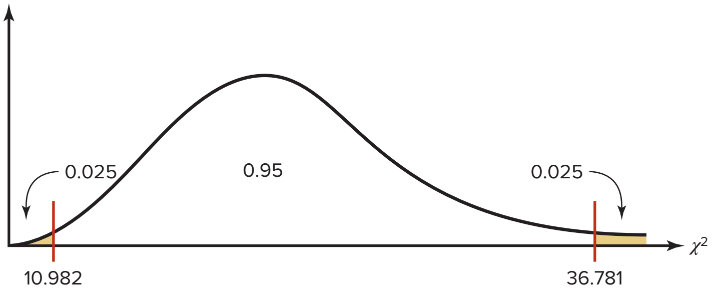
Find the critical chi-square values for 22 degrees of freedom when $\alpha = 0.05$ and a two-tailed test is conducted.

SOLUTION

When a two-tailed test is conducted, the area must be split, as shown in Figure 8–32. Note that the area to the right of the larger value is 0.025 (0.05/2 or $\alpha/2$), and the area to the right of the smaller value is 0.975 (1.00 − 0.05/2 or 1 − $\alpha/2$).

FIGURE 8–32

Chi-Square Distribution for Example 8–23



Remember that chi-square values cannot be negative. Hence, you must use α values in the table of 0.025 and 0.975. With 22 degrees of freedom, the critical values are 36.781 and 10.982, respectively.

Table G gives values only up to 30 degrees of freedom. When the degrees of freedom exceed those specified in the table, use the table value for 30 degrees of freedom. This guideline keeps the type 1 error equal to or below the α value.

When you are testing a claim about a single variance using the **chi-square test,** there are three possible test situations: right-tailed test, left-tailed test, and two-tailed test.

If a researcher believes the variance of a population to be greater than some specific value, say, 225, then the researcher states the hypotheses as

$$H_0\colon \sigma^2 = 225 \qquad \text{and} \qquad H_1\colon \sigma^2 > 225$$

and conducts a right-tailed test.

If the researcher believes the variance of a population to be less than 225, then the researcher states the hypotheses as

$$H_0\colon \sigma^2 = 225 \qquad \text{and} \qquad H_1\colon \sigma^2 < 225$$

and conducts a left-tailed test.

Finally, if a researcher does not wish to specify a direction, she or he states the hypotheses as

$$H_0\colon \sigma^2 = 225 \qquad \text{and} \qquad H_1\colon \sigma^2 \neq 225$$

and conducts a two-tailed test.

Formula for the Chi-Square Test for a Single Variance

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

with degrees of freedom equal to $n - 1$ and where

n = sample size
s^2 = sample variance
σ^2 = population variance

Unusual Stat

About 55% of cats owned in the United States are overweight.

Three assumptions are made for the chi-square test, as outlined here.

Assumptions for the Chi-Square Test for a Single Variance

1. The sample must be randomly selected from the population.
2. The population must be normally distributed for the variable under study.
3. The observations must be independent of one another.

In this book, the assumptions will be stated in the exercises; however, when encountering statistics in other situations, you must check to see that these assumptions have been met before proceeding.

The methods used to test a hypothesis about a population variance or population standard deviation are not robust. So if the data come from a population that is not normally distributed, these methods cannot be used.

The traditional method for hypothesis testing follows the same five steps listed earlier. They are repeated here.

Step 1 State the hypotheses and identify the claim.

Step 2 Find the critical value(s).

Step 3 Compute the test value.

Step 4 Make the decision.

Step 5 Summarize the results.

Examples 8–24 through 8–26 illustrate the traditional hypothesis-testing procedure for variances.

You might ask, Why is it important to test variances? There are several reasons. First, in any situation where consistency is required, such as in manufacturing, you would like to have the smallest variation possible in the products. For example, when bolts are manufactured, the variation in diameters due to the process must be kept to a minimum, or else the nuts will not fit them properly. In education, consistency is required on a test. That is, if the same students take the same test several times, they should get approximately the same grades, and the variance of each of the student's grades should be small. On the other hand, if the test is to be used to judge learning, the overall standard deviation of all the grades should be large so that you can differentiate those who have learned the subject from those who have not learned it.

EXAMPLE 8–24 IQ Test

A psychologist wishes to see if the variance in IQ of 10 of her counseling patients is less than the variance of the population, which is 225. The variance of the IQs of her 10 patients was 206. Test her claim at $\alpha = 0.05$.

SOLUTION

Step 1 State the hypotheses.

$$H_0\colon \sigma^2 = 225 \qquad \text{and} \qquad H_1\colon \sigma^2 < 225 \text{ (claim)}$$

Step 2 Find the critical value. Since this is a left-tailed test and $\alpha = 0.05$, the value $1 - 0.05 = 0.95$ should be used. The degrees of freedom are $n - 1 = 10 - 1 = 9$. Hence, the critical value is 3.325.

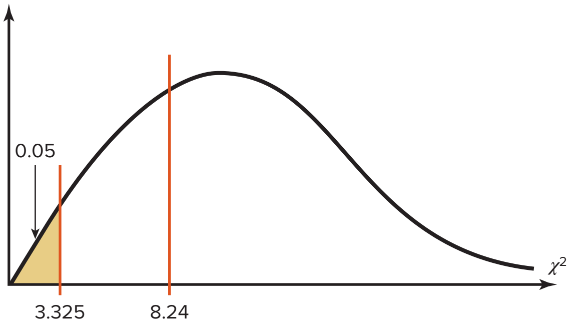
Step 3 Compute the test value.

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(10-1)(206)}{225} = 8.24$$

Step 4 Make the decision. Since 8.24 falls in the noncritical region, do not reject the null hypothesis. See Figure 8–33.

FIGURE 8–33

Critical and Test Value for Example 8–33



Step 5 Summarize the results. There is not enough evidence to support the claim that the variance of her 10 patients is less than 225.

EXAMPLE 8–25 Outpatient Surgery

A hospital administrator believes that the standard deviation of the number of people using outpatient surgery per day is greater than 8. A random sample of 15 days is selected. The data are shown. At $\alpha = 0.10$, is there enough evidence to support the administrator's claim? Assume the variable is normally distributed.

25	30	5	15	18
42	16	9	10	12
12	38	8	14	27

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon \sigma = 8 \quad \text{and} \quad H_1\colon \sigma > 8 \text{ (claim)}$$

Since the standard deviation is given, it should be squared to get the variance.

Step 2 Find the critical value. Since this test is right-tailed with d.f. of 15 − 1 = 14 and $\alpha = 0.10$, the critical value is 21.064.

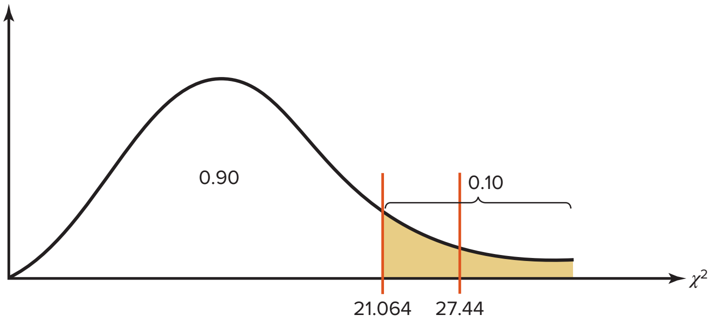
Step 3 Compute the test value. Since raw data are given, the standard deviation of the sample must be found by using the formula in Chapter 3 or your calculator. It is $s = 11.2$.

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(15-1)(11.2)^2}{64} = 27.44$$

Step 4 Make the decision. The decision is to reject the null hypothesis since the test value, 27.44, is greater than the critical value, 21.064, and falls in the critical region. See Figure 8–34.

FIGURE 8–34

Critical and Test Value for Example 8–25



Step 5 Summarize the results. There is enough evidence to support the claim that the standard deviation is greater than 8.

EXAMPLE 8–26 Nicotine Content of Cigarettes

A cigarette manufacturer wishes to test the claim that the variance of the nicotine content of its cigarettes is 0.644. Nicotine content is measured in milligrams, and assume that it is normally distributed. A random sample of 20 cigarettes has a standard deviation of 1.00 milligram. At $\alpha = 0.05$, is there enough evidence to reject the manufacturer's claim?

SOLUTION

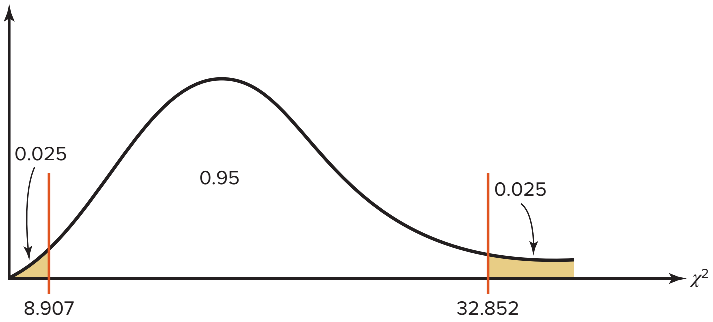
Step 1 State the hypotheses and identify the claim.

$$H_0\colon \sigma^2 = 0.644 \text{ (claim)} \qquad \text{and} \qquad H_1\colon \sigma^2 \neq 0.644$$

Step 2 Find the critical values. Since this test is a two-tailed test at $\alpha = 0.05$, the critical values for 0.025 and 0.975 must be found. The degrees of freedom are 19; hence, the critical values are 32.852 and 8.907, respectively. The critical or rejection regions are shown in Figure 8–35.

FIGURE 8–35

Critical Values for Example 8–26



Step 3 Compute the test value.

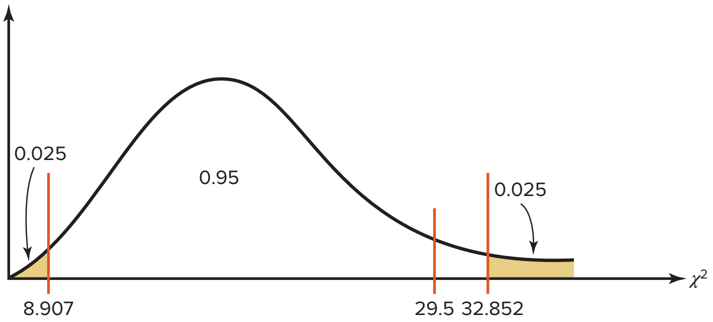
$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(20-1)(1.0)^2}{0.644} = 29.5$$

Since the sample standard deviation s is given in the problem, it must be squared for the formula.

Step 4 Make the decision. Do not reject the null hypothesis, since the test value falls between the critical values ($8.907 < 29.5 < 32.852$) and in the noncritical region, as shown in Figure 8–36.

FIGURE 8–36

Critical and Test Values for Example 8–26



Step 5 Summarize the results. There is not enough evidence to reject the manufacturer's claim that the variance of the nicotine content of the cigarettes is equal to 0.644.

Approximate P-values for the chi-square test can be found by using Table G in Appendix A. The procedure is somewhat more complicated than the previous procedures for finding *P*-values for the *z* and *t* tests since the chi-square distribution is not exactly symmetric and χ^2 values cannot be negative. As we did for the *t* test, we will determine an *interval* for the *P*-value based on the table. Examples 8–27 through 8–29 show the procedure.

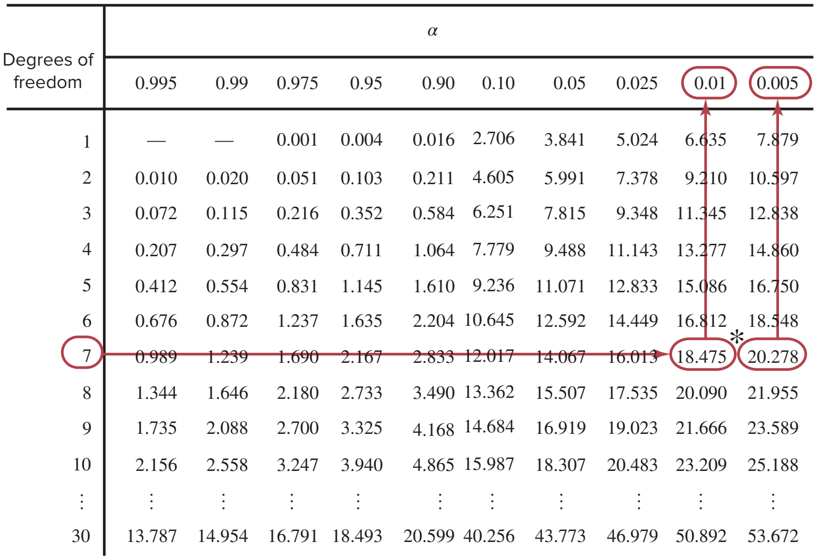
EXAMPLE 8–27

Find the *P*-value when $\chi^2 = 19.274$, $n = 8$, and the test is right-tailed.

SOLUTION

To get the *P*-value, look across the row with d.f. = 7 in Table G and find the two values that 19.274 falls between. They are 18.475 and 20.278. Look up to the top row and find the α values corresponding to 18.475 and 20.278. They are 0.01 and 0.005, respectively. See Figure 8–37. Hence, the *P*-value is contained in the interval $0.005 < P\text{-value} < 0.01$. (The *P*-value obtained from a calculator is 0.007.)

FIGURE 8–37 *P*-Value Interval for Example 8–27



Degrees of freedom	α									
	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	—	—	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.071	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
30	13.787	14.954	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672

*19.274 falls between 18.475 and 20.278

EXAMPLE 8–28

Find the *P*-value when $\chi^2 = 3.823$, $n = 13$, and the test is left-tailed.

SOLUTION

To get the *P*-value, look across the row with d.f. = 12 and find the two values that 3.823 falls between. They are 3.571 and 4.404. Look up to the top row and find the values corresponding to 3.571 and 4.404. They are 0.99 and 0.975, respectively. When the χ^2 test value falls on the left side, each of the values must be subtracted from 1 to get the interval that *P*-value falls between.

$$1 - 0.99 = 0.01 \qquad \text{and} \qquad 1 - 0.975 = 0.025$$

Hence, the *P*-value falls in the interval

$$0.01 < P\text{-value} < 0.025$$

(The *P*-value obtained from a calculator is 0.014.)

When the χ^2 test is two-tailed, both interval values must be doubled. If a two-tailed test were being used in Example 8–28, then the interval would be $2(0.01) < P\text{-value} < 2(0.025)$, or $0.02 < P\text{-value} < 0.05$.

The *P*-value method for hypothesis testing for a variance or standard deviation follows the same steps shown in the preceding sections.

Step 1 State the hypotheses and identify the claim.

Step 2 Compute the test value.

Step 3 Find the *P*-value.

Step 4 Make the decision.

Step 5 Summarize the results.

Example 8–29 shows the *P*-value method for variances or standard deviations.

EXAMPLE 8–29 Car Inspection Times

A researcher knows from past studies that the standard deviation of the time it takes to inspect a car is 16.8 minutes. A random sample of 24 cars is selected and inspected. The standard deviation is 12.5 minutes. At $\alpha = 0.05$, can it be concluded that the standard deviation has changed? Use the *P*-value method. Assume the variable is normally distributed.

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon \sigma = 16.8 \qquad \text{and} \qquad H_1\colon \sigma \neq 16.8 \text{ (claim)}$$

Step 2 Compute the test value.

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(24-1)(12.5)^2}{(16.8)^2} = 12.733$$

Step 3 Find the *P*-value. Using Table G with d.f. = 23, the value 12.733 falls between 11.689 and 13.091, corresponding to 0.975 and 0.95, respectively. Since these values are found on the left side of the distribution, each value must be subtracted from 1. Hence, $1 - 0.975 = 0.025$ and $1 - 0.95 = 0.05$. Since this is a two-tailed test, the area must be doubled to obtain the *P*-value interval. Hence, $0.05 < P\text{-value} < 0.10$, or somewhere between 0.05 and 0.10. (The *P*-value obtained from a calculator is 0.085.)

Step 4 Make the decision. Since $\alpha = 0.05$ and the *P*-value is between 0.05 and 0.10, the decision is to not reject the null hypothesis since $P\text{-value} > \alpha$.

Step 5 Summarize the results. There is not enough evidence to support the claim that the standard deviation of the time it takes to inspect a car has changed.

Applying the Concepts 8–5

Testing Gas Mileage Claims

Assume that you are working for the Consumer Protection Agency and have recently been getting complaints about the highway gas mileage of the new Dodge Caravans. Chrysler Corporation agrees to allow you to randomly select 40 of its new Dodge Caravans to test the highway mileage.

Chrysler claims that the Caravans get 28 mpg on the highway. Your results show a mean of 26.7 and a standard deviation of 4.2. You support Chrysler's claim.

1. Show whether or not you support Chrysler's claim by listing the *P*-value from your output. After more complaints, you decide to test the variability of the miles per gallon on the highway. From further questioning of Chrysler's quality control engineers, you find they are claiming a standard deviation of no more than 2.1. Use a one-tailed test.
2. Test the claim about the standard deviation.
3. Write a short summary of your results and any necessary action that Chrysler must take to remedy customer complaints.
4. State your position about the necessity to perform tests of variability along with tests of the means.

See page 486 for the answers.

Exercises 8–5

1. Using Table G, find the critical value(s) for each. Indicate the noncritical region or regions, and state the null and alternative hypotheses. Use $\sigma^2 = 225$.

a. $\alpha = 0.10$, $n = 14$, two-tailed
b. $\alpha = 0.05$, $n = 27$, right-tailed
c. $\alpha = 0.01$, $n = 9$, left-tailed
d. $\alpha = 0.05$, $n = 17$, right-tailed

2. Using Table G, find the critical value(s) for each. Show the critical and noncritical regions, and state the appropriate null and alternative hypotheses. Use $\sigma^2 = 225$.

a. $\alpha = 0.01$, $n = 17$, right-tailed
b. $\alpha = 0.025$, $n = 20$, left-tailed
c. $\alpha = 0.01$, $n = 13$, two-tailed
d. $\alpha = 0.025$, $n = 29$, left-tailed

3. Using Table G, find the *P*-value interval for each χ^2 test value.

a. $\chi^2 = 29.321$, $n = 16$, right-tailed
b. $\chi^2 = 10.215$, $n = 25$, left-tailed
c. $\chi^2 = 24.672$, $n = 11$, two-tailed
d. $\chi^2 = 23.722$, $n = 9$, right-tailed

4. Using Table G, find the *P*-value interval for each χ^2 test value.

a. $\chi^2 = 13.974$, $n = 28$, two-tailed
b. $\chi^2 = 10.571$, $n = 19$, left-tailed
c. $\chi^2 = 12.144$, $n = 6$, two-tailed
d. $\chi^2 = 8.201$, $n = 23$, two-tailed

For Exercises 5 through 20, assume that the variables are normally or approximately normally distributed. Use the traditional method of hypothesis testing unless otherwise specified.

5. Age of Psychologists Test the claim that the standard deviation of the ages of psychologists in Pennsylvania is 8.6 at $\alpha = 0.05$. A random sample of 12 psychologists had a standard deviation of 9.3.

6. Carbohydrates in Fast Foods The number of carbohydrates found in a random sample of fast-food entrees is listed. Is there sufficient evidence to conclude that the variance differs from 100? Use the 0.05 level of significance.

53	46	39	39	30
47	38	73	43	41

Source: Fast Food Explorer (www.fatcalories.com).

7. Transferring Phone Calls The manager of a large company claims that the standard deviation of the time (in minutes) that it takes a telephone call to be transferred to the correct office in her company is 1.2 minutes or less. A random sample of 15 calls is selected, and the calls are timed. The standard deviation of the sample is 1.8 minutes. At $\alpha = 0.01$, test the claim that the standard deviation is less than or equal to 1.2 minutes. Use the *P*-value method.

8. Soda Bottle Content A machine fills 12-ounce bottles with soda. For the machine to function properly, the standard deviation of the population must be less than or equal to 0.03 ounce. A random sample of 8 bottles is selected, and the number of ounces of soda in each bottle is given. At $\alpha = 0.05$, can we reject the claim that the machine is functioning properly? Use the *P*-value method.

12.03	12.10	12.02	11.98
12.00	12.05	11.97	11.99

9. **Distances to Supermarkets** A random sample of the distances in miles 8 shoppers travel to their nearest supermarkets is shown. Test the claim at $\alpha = 0.10$ that the standard deviation of the distance shoppers travel is greater than 2 miles.

3.6 4.2 1.7 1.3
5.1 9.3 2.9 6.5

10. **Exam Grades** A statistics professor is used to having a variance in his class grades of no more than 100. He feels that his current group of students is different, and so he examines a random sample of midterm grades as shown. At $\alpha = 0.05$, can it be concluded that the variance in grades exceeds 100?

92.3 89.4 76.9 65.2 49.1
96.7 69.5 72.8 67.5 52.8
88.5 79.2 72.9 68.7 75.8

11. **Tornado Deaths** A researcher claims that the standard deviation of the number of deaths annually from tornadoes in the United States is less than 35. If a random sample of 11 years had a standard deviation of 32, is the claim believable? Use $\alpha = 0.05$.

Source: National Oceanic and Atmospheric Administration.

12. **Interstate Speeds** It has been reported that the standard deviation of the speeds of drivers on Interstate 75 near Findlay, Ohio, is 8 miles per hour for all vehicles. A driver feels from experience that this is very low. A survey is conducted, and for 50 randomly selected drivers the standard deviation is 10.5 miles per hour. At $\alpha = 0.05$, is the driver correct?

13. **Nicotine Content of Cigarettes** A manufacturer of cigarettes wishes to test the claim that the variance of the nicotine content of the cigarettes the company manufactures is equal to 0.638 milligram. The variance of a random sample of 25 cigarettes is 0.930 milligram. At $\alpha = 0.05$, test the claim.

14. **Vitamin C in Fruits and Vegetables** The amounts of vitamin C (in milligrams) for 100 g (3.57 ounces) of various randomly selected fruits and vegetables are listed. Is there sufficient evidence to conclude that the standard deviation differs from 12 mg? Use $\alpha = 0.10$.

7.9 16.3 12.8 13.0 32.2 28.1 34.4
46.4 53.0 15.4 18.2 25.0 5.2

Source: Time Almanac 2012.

15. **Manufactured Machine Parts** A manufacturing process produces machine parts with measurements the standard deviation of which must be no more than 0.52 mm. A random sample of 20 parts in a given lot revealed a standard deviation in measurement of 0.568 mm. Is there sufficient evidence at $\alpha = 0.05$ to conclude that the standard deviation of the parts is outside the required guidelines?

16. **Golf Scores** A random sample of second-round golf scores from a major tournament is listed below. At $\alpha = 0.10$, is there sufficient evidence to conclude that the population variance exceeds 9?

75 67 69 72 70
66 74 69 74 71

17. **Calories in Pancake Syrup** A nutritionist claims that the standard deviation of the number of calories in 1 tablespoon of the major brands of pancake syrup is 60. A random sample of major brands of syrup is selected, and the number of calories is shown. At $\alpha = 0.10$, can the claim be rejected?

53 210 100 200 100 220
210 100 240 200 100 210
100 210 100 210 100 60

Source: Based on information from *The Complete Book of Food Counts* by Corrine T. Netzer, Dell Publishers, New York.

18. **High Temperatures in January** Daily weather observations for southwestern Pennsylvania for the first three weeks of January for randomly selected years show daily high temperatures as follows: 55, 44, 51, 59, 62, 60, 46, 51, 37, 30, 46, 51, 53, 57, 57, 39, 28, 37, 35, and 28 degrees Fahrenheit. The normal standard deviation in high temperatures for this time period is usually no more than 8 degrees. A meteorologist believes that with the unusual trend in temperatures the standard deviation is greater. At $\alpha = 0.05$, can we conclude that the standard deviation is greater than 8 degrees?

Source: www.wunderground.com

19. **College Room and Board Costs** Room and board fees for a random sample of independent religious colleges are shown.

7460 7959 7650 8120 7220
8768 7650 8400 7860 6782
8754 7443 9500 9100

Estimate the standard deviation in costs based on $s \approx R/4$. Is there sufficient evidence to conclude that the sample standard deviation differs from this estimated amount? Use $\alpha = 0.05$.

Source: World Almanac.

20. **Heights of Volcanoes** A random sample of heights (in feet) of active volcanoes in North America, outside of Alaska, is shown. Is there sufficient evidence that the standard deviation in heights of volcanoes outside Alaska is less than the standard deviation in heights of Alaskan volcanoes, which is 2385.9 feet? Use $\alpha = 0.05$.

10,777 8159 11,240 10,456
14,163 8363

Source: Time Almanac.

Technology Step by Step

TI-84 Plus
Step by Step

The TI-84 Plus does not have a built-in hypothesis test for the variance or standard deviation. However, the downloadable program named SDHYP is available in your online resources. Follow the instructions online for downloading the program.

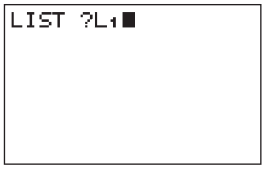
Performing a Hypothesis Test for the Variance and Standard Deviation (Data)

1. Enter the values into L_1.
2. Press **PRGM,** move the cursor to the program named SDHYP, and press **ENTER** twice.
3. Press **1** for Data.
4. Type **L_1** for the list and press **ENTER.**
5. Type the number corresponding to the type of alternative hypothesis.
6. Type the value of the hypothesized variance and press **ENTER.**
7. Press **ENTER** to clear the screen.

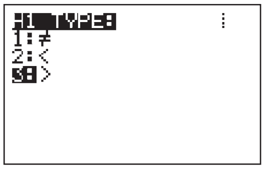
Example TI8–4

This pertains to Example 8–25 in the text. Test the claim that $\sigma > 8$ for these data.

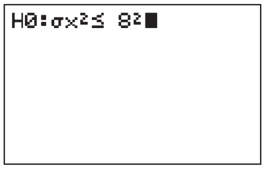
25 30 5 15 18 42 16 9 10 12 12 38 8 14 27



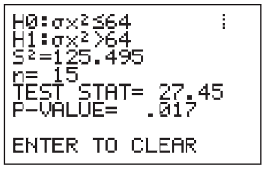
```
LIST ?L₁■
```



```
H1 TYPE?
1:≠
2:<
3:>
```



```
H0:σx²≤ 8²■
```



```
H0:σx²≤64
H1:σx²>64
S²=125.495
n= 15
TEST STAT= 27.45
P-VALUE=  .017

ENTER TO CLEAR
```

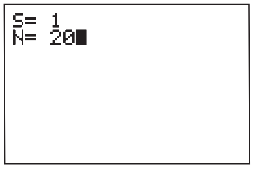
Since P-value $= 0.017 < 0.1$, we reject H_0 and conclude H_1. Therefore, there is enough evidence to support the claim that the standard deviation of the number of people using outpatient surgery is greater than 8.

Performing a Hypothesis Test for the Variance and Standard Deviation (Statistics)

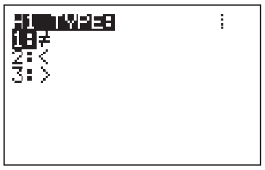
1. Press **PRGM,** move the cursor to the program named SDHYP, and press **ENTER** twice.
2. Press **2** for Stats.
3. Type the sample standard deviation and press **ENTER.**
4. Type the sample size and press **ENTER.**
5. Type the number corresponding to the type of alternative hypothesis.
6. Type the value of the hypothesized variance and press **ENTER.**
7. Press **ENTER** to clear the screen.

Example TI8–5

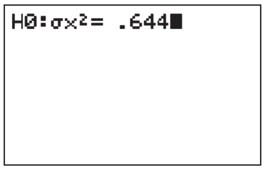
This pertains to Example 8–26 in the text. Test the claim that $\sigma^2 = 0.644$, given $n = 20$ and $s = 1$.



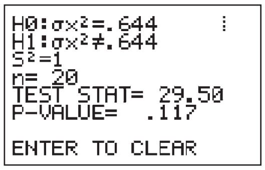
```
S= 1
N= 20■
```



```
H1 TYPE?
1:≠
2:<
3:>
```



```
H0:σx²= .644■
```



```
H0:σx²=.644
H1:σx²≠.644
S²=1
n= 20
TEST STAT= 29.50
P-VALUE=  .117

ENTER TO CLEAR
```

Since P-value $= 0.117 > 0.05$, we do not reject H_0 and do not conclude H_1. Therefore, there is not enough evidence to reject the manufacturer's claim that the variance of the nicotine content of the cigarettes is equal to 0.644.

EXCEL

Step by Step

Hypothesis Test for the Variance: Chi-Square Test

Excel does not have a procedure to conduct a hypothesis test for a single population variance. However, you may conduct the test of the variance using the MegaStat Add-in available in your online resources. If you have not installed this add-in, do so, following the instructions from the Chapter 1 Excel Step by Step.

Example XL8–3

This example relates to Example 8–26 from the text. At the 5% significance level, test the claim that $\sigma^2 = 0.644$. The MegaStat chi-square test of the population variance uses the P-value method. Therefore, it is not necessary to enter a significance level.

1. Type a label for the variable: Nicotine in cell A1.
2. Type the observed variance: **1** in cell A2.
3. Type the sample size: **20** in cell A3.
4. From the toolbar, select Add-Ins, **MegaStat>Hypothesis Tests>Chi-Square Variance Test.** *Note:* You may need to open MegaStat from the MegaStat.xls file on your computer's hard drive.
5. Select summary input.
6. Type **A1:A3** for the Input Range.
7. Type **0.644** for the Hypothesized variance and select the Alternative not equal.
8. Click [OK].

The result of the procedure is shown next.

```
Chi-Square Variance Test
  0.64   Hypothesized variance
  1.00   Observed variance of nicotine
    20   n
    19   d.f.
 29.50   Chi-square

0.1169   P-value (two-tailed)
```

MINITAB

Step by Step

Hypothesis Test for Standard Deviation or Variance

MINITAB can be used to find a critical value of chi-square. It can also calculate the test statistic and P-value for a chi-square test of variance.

Example 8–22

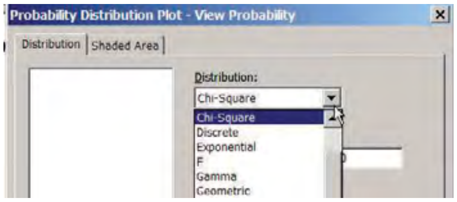
Find the critical χ^2 value for $\alpha = 0.05$ for a left-tailed test with d.f. = 10.

Step 1 To find the critical value of t for a right-tailed test, select **Graph>Probability Distribution Plot,** then **View Probability,** then click [**OK**].

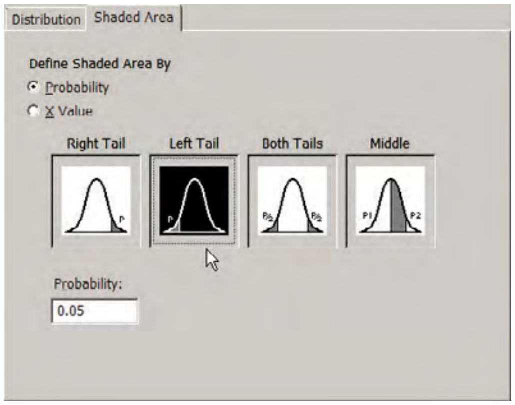
Step 2 Change the Distribution to a Chi-square distribution and type in the degrees of freedom, **10.**

Step 3 Click the tab for **Shaded Area.**

a) Select the ratio button for **Probability.**
b) Select **Left Tail.**
c) Type in the value of alpha for probability, **0.05.**
d) Click [**OK**].

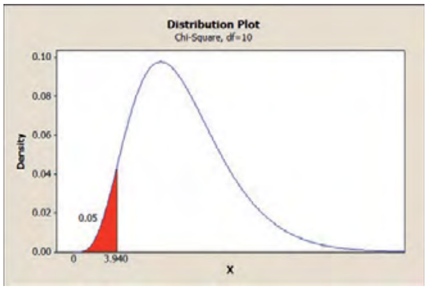

Probability Distribution Plot - View Probability

Distribution: Chi-Square


Define Shaded Area By
● Probability
○ X Value

Left Tail

Probability: 0.05



The critical value of χ^2 to three decimal places is 3.940.
You may click the Edit Last Dialog button and then change the settings for additional critical values.

Example 8–25 Outpatient Surgery

MINITAB will calculate the test statistic and *P*-value. There are data for this example.

Step 1 Type the data into a new MINITAB worksheet. All 15 values must be in C1. Type the label Surgeries above the first row of data.

Step 2 Select **Stat>Basic Statistics> 1-variance.**

Step 3 Select **One or more samples, each in a column** from the drop down menu, if not already selected.

Step 4 To select the data, click inside the next dialog box for Columns; then select C1 Surgeries from the list.

Step 5 Select the box for Perform hypothesis test.

a) Select Hypothesized standard deviation from the drop-down list.

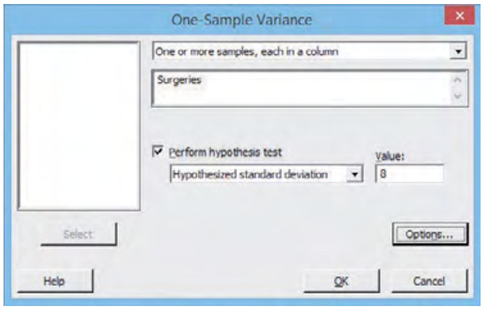
b) Type in the hypothesized value of **8.**

Step 6 Click the button for **[Options]**.

a) Type the default confidence level, that is, **90.**

b) Click the drop-down menu for the Alternative hypothesis, greater than.

Step 7 Click **[OK]** twice.


One-Sample Variance

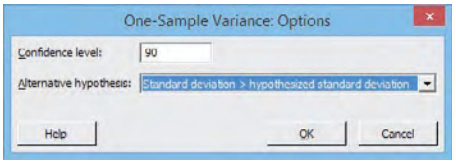
One or more samples, each in a column

Surgeries

☑ Perform hypothesis test

Hypothesized standard deviation

Value: 8


One-Sample Variance: Options

Confidence level: 90

Alternative hypothesis: Standard deviation > hypothesized standard deviation

In the Session Window scroll down to the output labeled Statistics and further to the output labeled Tests. You should see the test statistic and *P*-value for the chi-square test. Since the *P*-value is less than 0.10, the null hypothesis will be rejected. The standard deviation, $s = 11.2$ is significantly greater than 8.

```
Statistics

Variable     N      StDev    Variance
Surgeries    15      11.2         125

Tests

                             Test
Variable     Method       Statistic      DF      P-Value
Surgeries    Chi-Square       27.45      14        0.017
```

Although the text shows how to calculate a *P*-value, these are included in the MINITAB output of all hypothesis tests. The Alternative hypothesis in the Options dialog box must match your Alternative hypothesis.

8–6 Additional Topics Regarding Hypothesis Testing

In hypothesis testing, there are several other concepts that might be of interest to students in elementary statistics. These topics include the relationship between hypothesis testing and confidence intervals, and some additional information about the type II error.

OBJECTIVE 9

Test hypotheses, using confidence intervals.

Confidence Intervals and Hypothesis Testing

There is a relationship between confidence intervals and hypothesis testing. When the null hypothesis is rejected in a hypothesis-testing situation, the confidence interval for the mean using the same level of significance *will not* contain the hypothesized mean. Likewise, when the null hypothesis is not rejected, the confidence interval computed using the same level of significance *will* contain the hypothesized mean. Examples 8–30 and 8–31 show this concept for two-tailed tests.

EXAMPLE 8–30 Sugar Packaging

Sugar is packed in 5-pound bags. An inspector suspects the bags may not contain 5 pounds. A random sample of 50 bags produces a mean of 4.6 pounds and a standard deviation of 0.7 pound. Is there enough evidence to conclude that the bags do not contain 5 pounds as stated at $\alpha = 0.05$? Also, find the 95% confidence interval of the true mean. Assume the variable is normally distributed.

© Nancy R. Cohen/Getty Images RF

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon \mu = 5 \qquad \text{and} \qquad H_1\colon \mu \neq 5 \text{ (claim)}$$

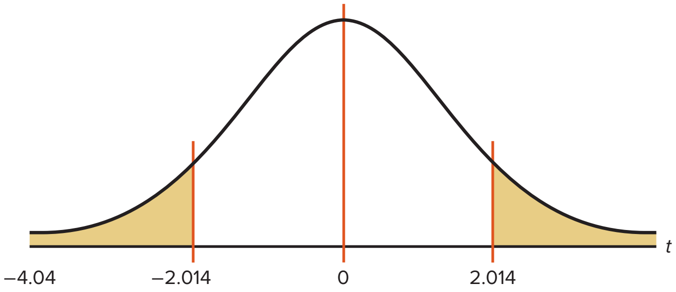
Step 2 At $\alpha = 0.05$ and d.f. = 49 (use d.f. = 45), the critical values are +2.014 and −2.014.

Step 3 Compute the test value.

$$t = \frac{\overline{X} - \mu}{s/\sqrt{n}} = \frac{4.6 - 5.0}{0.7/\sqrt{50}} = \frac{-0.4}{0.099} = -4.04$$

Step 4 Make the decision. Reject the null hypothesis since $-4.04 < -2.014$. See Figure 8–38.

FIGURE 8–38
Critical Values and Test Value for Example 8–30



Step 5 Summarize the results. There is enough evidence to support the claim that the bags do not weigh 5 pounds.

The 95% confidence for the mean is given by

$$\overline{X} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \overline{X} + t_{\alpha/2}\frac{s}{\sqrt{n}}$$

$$4.6 - (2.014)\left(\frac{0.7}{\sqrt{50}}\right) < \mu < 4.6 + (2.014)\left(\frac{0.7}{\sqrt{50}}\right)$$

$$4.4 < \mu < 4.8$$

Notice that the 95% confidence interval of μ does *not* contain the hypothesized value $\mu = 5$. Hence, there is agreement between the hypothesis test and the confidence interval.

EXAMPLE 8–31 Hog Weights

A researcher claims that adult hogs fed a special diet will have an average weight of 200 pounds. A random sample of 10 hogs has an average weight of 198.2 pounds and a standard deviation of 3.3 pounds. At $\alpha = 0.05$, can the claim be rejected? Also, find the 95% confidence interval of the true mean. Assume the variable is normally distributed.

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon \mu = 200 \text{ lb (claim)} \qquad \text{and} \qquad H_1\colon \mu \neq 200 \text{ lb}$$

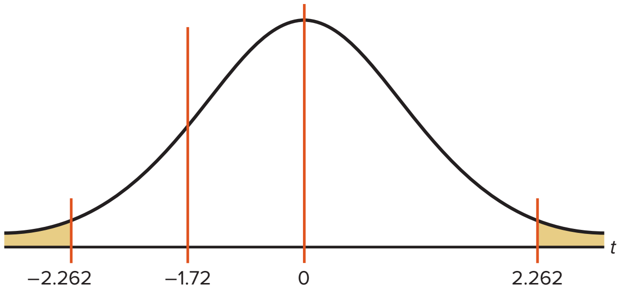
Step 2 Find the critical values. At $\alpha = 0.05$ and d.f. = 9, the critical values are +2.262 and −2.262.

Step 3 Compute the test value.

$$t = \frac{\overline{X} - \mu}{s/\sqrt{n}} = \frac{198.2 - 200}{3.3/\sqrt{10}} = \frac{-1.8}{1.0436} = -1.72$$

Step 4 Make the decision. Do not reject the null hypothesis. See Figure 8–39.

FIGURE 8–39
Critical Values and Test Value for Example 8–31



Step 5 Summarize the results. There is not enough evidence to reject the claim that the mean weight of adult hogs is 200 lb.

The 95% confidence interval of the mean is

$$\overline{X} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \overline{X} + t_{\alpha/2}\frac{s}{\sqrt{n}}$$

$$198.2 - (2.262)\left(\frac{3.3}{\sqrt{10}}\right) < \mu < 198.2 + (2.262)\left(\frac{3.3}{\sqrt{10}}\right)$$

$$198.2 - 2.361 < \mu < 198.2 + 2.361$$

$$195.8 < \mu < 200.6$$

The 95% confidence interval does contain the hypothesized mean $\mu = 200$. Again there is agreement between the hypothesis test and the confidence interval.

In summary, then, when the null hypothesis is rejected at a significance level of α, the confidence interval computed at the $1 - \alpha$ level will not contain the value of the mean that is stated in the null hypothesis. On the other hand, when the null hypothesis is not rejected, the confidence interval computed at the same significance level will contain the value of the mean stated in the null hypothesis. These results are true for other hypothesis-testing situations and are not limited to means tests.

The relationship between confidence intervals and hypothesis testing presented here is valid for two-tailed tests. The relationship between one-tailed hypothesis tests and one-sided or one-tailed confidence intervals is also valid; however, this technique is beyond the scope of this text.

Type II Error and the Power of a Test

Recall that in hypothesis testing, there are two possibilities: Either the null hypothesis H_0 is true, or it is false. Furthermore, on the basis of the statistical test, the null hypothesis is either rejected or not rejected. These results give rise to four possibilities, as shown in Figure 8–40. This figure is similar to Figure 8–2.

OBJECTIVE 10

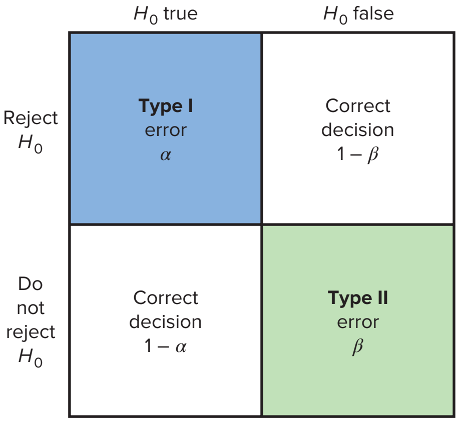
Explain the relationship between type I and type II errors and the power of a test.

As stated previously, there are two types of errors: type I and type II. A type I error can occur only when the null hypothesis is rejected. By choosing a level of significance, say, of 0.05 or 0.01, the researcher can determine the probability of committing a type I error. For example, suppose that the null hypothesis was H_0: $\mu = 50$, and it was rejected. At the 0.05 level (one tail), the researcher has only a 5% chance of being wrong, i.e., of rejecting a true null hypothesis.

On the other hand, if the null hypothesis is not rejected, then either it is true or a type II error has been committed. A type II error occurs when the null hypothesis is indeed false, but is not rejected. The probability of committing a type II error is denoted as β.

The value of β is not easy to compute. It depends on several things, including the value of α, the size of the sample, the population standard deviation, and the actual difference between the hypothesized value of the parameter being tested and the true parameter. The researcher has control over two of these factors, namely, the selection of α and the size of the sample. The standard deviation of the population is sometimes known or can be estimated. The major problem, then, lies in knowing the actual difference between the hypothesized parameter and the true parameter. If this difference were known, then the value of the parameter would be known; and if the parameter were known, then there would be no need to do any hypothesis testing. Hence, the value of β cannot be computed. But this does not mean that it should be ignored. What the researcher usually does is to try to minimize the size of β or to maximize the size of $1 - \beta$, which is called the **power of a test.**

FIGURE 8–40
Possibilities in Hypothesis Testing



	H_0 true	H_0 false
Reject H_0	**Type I** error α	Correct decision $1 - \beta$
Do not reject H_0	Correct decision $1 - \alpha$	**Type II** error β

The power of a statistical test measures the sensitivity of the test to detect a real difference in parameters if one actually exists. The power of a test is a probability and, like all probabilities, can have values ranging from 0 to 1. The higher the power, the more sensitive the test is to detecting a real difference between parameters if there is a difference. In other words, the closer the power of a test is to 1, the better the test is for rejecting the null hypothesis if the null hypothesis is, in fact, false.

The power of a test is equal to $1 - \beta$, that is, 1 minus the probability of committing a type II error. The power of the test is shown in the upper right-hand block of Figure 8–40. If somehow it were known that $\beta = 0.04$, then the power of a test would be $1 - 0.04 = 0.96$, or 96%. In this case, the probability of rejecting the null hypothesis when it is false is 96%.

As stated previously, the power of a test depends on the probability of committing a type II error, and since β is not easily computed, the power of a test cannot be easily computed. (See the Critical Thinking Challenges on pages 484 and 485.)

However, there are some guidelines that can be used when you are conducting a statistical study concerning the power of a test. In that case, use the test that has the highest power for the data. There are times when the researcher has a choice of two or more statistical tests to test the hypotheses. The tests with the highest power should be used. It is important, however, to remember that statistical tests have assumptions that need to be considered.

If these assumptions cannot be met, then another test with lower power should be used. The power of a test can be increased by increasing the value of α. For example, instead of using $\alpha = 0.01$, use $\alpha = 0.05$. Recall that as α increases, β decreases. So if β is decreased, then $1 - \beta$ will increase, thus increasing the power of the test.

Another way to increase the power of a test is to select a larger sample size. A larger sample size would make the standard error of the mean smaller and consequently reduce β. (The derivation is omitted.)

These two methods should not be used at the whim of the researcher. Before α can be increased, the researcher must consider the consequences of committing a type I error. If these consequences are more serious than the consequences of committing a type II error, then α should not be increased.

Likewise, there are consequences to increasing the sample size. These consequences might include an increase in the amount of money required to do the study and an increase in the time needed to tabulate the data. When these consequences result, increasing the sample size may not be practical.

There are several other methods a researcher can use to increase the power of a statistical test, but these methods are beyond the scope of this text.

One final comment is necessary. When the researcher fails to reject the null hypothesis, this does not mean that there is not enough evidence to support alternative hypotheses. It may be that the null hypothesis is false, but the statistical test has too low a power to detect the real difference; hence, one can conclude only that in this study, there is not enough evidence to reject the null hypothesis.

The relationship among α, β, and the power of a test can be analyzed in greater detail than the explanation given here. However, it is hoped that this explanation will show you that there is no magic formula or statistical test that can guarantee foolproof results when a decision is made about the validity of H_0. Whether the decision is to reject H_0 or not to reject H_0, there is in either case a chance of being wrong. The goal, then, is to try to keep the probabilities of type I and type II errors as small as possible.

Applying the Concepts 8–6

Consumer Protection Agency Complaints

Hypothesis testing and testing claims with confidence intervals are two different approaches that lead to the same conclusion. In the following activities, you will compare and contrast those two approaches.

Assume you are working for the Consumer Protection Agency and have recently been getting complaints about the highway gas mileage of the new Dodge Caravans. Chrysler Corporation agrees to allow you to randomly select 40 of its new Dodge Caravans to test the highway mileage. Chrysler claims that the vans get 28 mpg on the highway. Your results show a mean of 26.7 and a standard deviation of 4.2. You are not certain if you should create a confidence interval or run a hypothesis test. You decide to do both at the same time.

1. Draw a normal curve, labeling the critical values, critical regions, test statistic, and population mean. List the significance level and the null and alternative hypotheses.
2. Draw a confidence interval directly below the normal distribution, labeling the sample mean, error, and boundary values.
3. Explain which parts from each approach are the same and which parts are different.
4. Draw a picture of a normal curve and confidence interval where the sample and hypothesized means are equal.
5. Draw a picture of a normal curve and confidence interval where the lower boundary of the confidence interval is equal to the hypothesized mean.
6. Draw a picture of a normal curve and confidence interval where the sample mean falls in the left critical region of the normal curve.

See page 486 for the answers.

Exercises 8–6

1. **First-Time Births** According to the almanac, the mean age for a woman giving birth for the first time is 25.2 years. A random sample of ages of 35 professional women giving birth for the first time had a mean of 28.7 years and a standard deviation of 4.6 years. Use both a confidence interval and a hypothesis test at the 0.05 level of significance to test if the mean age of professional woman is different from 25.2 years at the time of their first birth.

2. **One-Way Airfares** The average one-way airfare from Pittsburgh to Washington, D.C., is $236. A random sample of 20 one-way fares during a particular month had a mean of $210 with a standard deviation of $43. At $\alpha = 0.02$, is there sufficient evidence to conclude a difference from the stated mean? Use the sample statistics to construct a 98% confidence interval for the true mean one-way airfare from Pittsburgh to Washington, D.C., and compare your interval to the results of the test. Do they support or contradict one another?

   Source: www.fedstats.gov

3. **IRS Audits** The IRS examined approximately 1% of individual tax returns for a specific year, and the average recommended additional tax per return was $19,150. Based on a random sample of 50 returns, the mean additional tax was $17,020. If the population standard deviation is $4080, is there sufficient evidence to conclude that the mean differs from $19,150 at $\alpha = 0.05$? Does a 95% confidence interval support this result?

   Source: New York Times Almanac.

4. **Prison Time** According to a public service website, 69.4% of white collar criminals get prison time. A randomly selected sample of 165 white collar criminals revealed that 120 were serving or had served prison time. Using $\alpha = 0.05$, test the conjecture that the proportion of white collar criminals serving prison time differs from 69.4% in two different ways.

5. **Working at Home** Workers with a formal arrangement with their employer to be paid for time worked at home worked an average of 19 hours per week. A random sample of 15 mortgage brokers indicated that they worked a mean of 21.3 hours per week at home with a standard deviation of 6.5 hours. At $\alpha = 0.05$, is there sufficient evidence to conclude a difference? Construct a 95% confidence interval for the true mean number of paid working hours at home. Compare the results of your confidence interval to the conclusion of your hypothesis test and discuss the implications.

   Source: www.bls.gov

6. **Newspaper Reading Times** A survey taken several years ago found that the average time a person spent reading the local daily newspaper was 10.8 minutes. The standard deviation of the population was 3 minutes. To see whether the average time had changed since the newspaper's format was revised, the newspaper editor surveyed 36 individuals. The average time that these 36 randomly selected people spent reading the paper was 12.2 minutes. At $\alpha = 0.02$, is there a change in the average time an individual spends reading the newspaper? Find the 98% confidence interval of the mean. Do the results agree? Explain.

7. What is meant by the power of a test?

8. How is the power of a test related to the type II error?

9. How can the power of a test be increased?

Summary

This chapter introduces the basic concepts of hypothesis testing. A statistical hypothesis is a conjecture about a population. There are two types of statistical hypotheses: the null and the alternative hypotheses. The null hypothesis states that there is no difference, and the alternative hypothesis specifies a difference. To test the null hypothesis, researchers use a statistical test. Many test values are computed by using

$$\text{Test value} = \frac{(\text{observed value}) - (\text{expected value})}{\text{standard error}}$$

- Researchers compute a test value from the sample data to decide whether the null hypothesis should be rejected. Statistical tests can be one-tailed or two-tailed, depending on the hypotheses.

  The null hypothesis is rejected when the difference between the population parameter and the sample statistic is said to be significant. The difference is significant when the test value falls in the critical region of the distribution. The critical region is determined by α, the level of significance of the test. The level is the probability of committing a type I error. This error occurs when the null hypothesis is rejected when it is true. Three generally agreed upon significance levels are 0.10, 0.05, and 0.01. A second kind of error, the type II error, can occur when the null hypothesis is not rejected when it is false. (8–1)
- There are two common methods used to test hypotheses; they are the traditional method and the *P*-value method. (8–2)
- All hypothesis-testing situations using the traditional method should include the following steps:
  1. State the null and alternative hypotheses and identify the claim.
  2. State an alpha level and find the critical value(s).
  3. Compute the test value.
  4. Make the decision to reject or not reject the null hypothesis.
  5. Summarize the results.
- All hypothesis-testing situations using the *P*-value method should include the following steps:
  1. State the hypotheses and identify the claim.
  2. Compute the test value.
  3. Find the *P*-value.
  4. Make the decision.
  5. Summarize the results.
- The z test is used to test a mean when the population standard deviation is known. When the sample size is less than 30, the population values need to be normally distributed. (8–2)
- When the population standard deviation is not known, researchers use a t test to test a claim about a mean. If the sample size is less than 30, the population values need to be normally or approximately normally distributed. (8–3)
- The z test can be used to test a claim about a population when $np \geq 5$ and $nq \geq 5$. (8–4)

- A single variance can be tested by using the chi-square test. (8–5)
- There is a relationship between confidence intervals and hypothesis testing. When the null hypothesis is rejected, the confidence interval for the mean using the same level of significance will not contain the hypothesized mean. When the null hypothesis is not rejected, the confidence interval, using the same level of significance, will contain the hypothesized mean. (8–6)
- The power of a statistical test measures the sensitivity of the test to detect a real difference in parameters if one actually exists; $1 - \beta$ is called the power of a test. (8–6)

Important Terms

α (alpha) 419

alternative hypothesis 414

β (beta) 419

chi-square test 463

critical or rejection region 420

critical value 420

hypothesis testing 414

left-tailed test 420

level of significance 419

noncritical or nonrejection region 420

null hypothesis 414

one-tailed test 420

power of a test 476

P-value 430

research hypothesis 415

right-tailed test 420

statistical hypothesis 414

statistical test 417

test value 417

t test 442

two-tailed test 421

type I error 418

type II error 418

z test 427

Important Formulas

Formula for the *z* test for means:

$$z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \quad \textbf{if } n < 30\textbf{, variable must be normally distributed}$$

Formula for the *t* test for means:

$$t = \frac{\overline{X} - \mu}{s/\sqrt{n}} \quad \textbf{if } n < 30\textbf{, variable must be normally distributed}$$

Formula for the *z* test for proportions:

$$z = \frac{\hat{p} - p}{\sqrt{pq/n}}$$

Formula for the chi-square test for variance or standard deviation:

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

Review Exercises

For Exercises 1 through 20, perform each of the following steps.

a. State the hypotheses and identify the claim.
b. Find the critical value(s).
c. Compute the test value.
d. Make the decision.
e. Summarize the results.

Use the traditional method of hypothesis testing unless otherwise specified.

Section 8–2

1. **Lifetime of \$1 Bills** The average lifetime of circulated \$1 bills is 18 months. A researcher believes that the average lifetime is not 18 months. He researched the lifetime of 50 \$1 bills and found the average lifetime was 18.8 months. The population standard deviation is 2.8 months. At $\alpha = 0.02$, can it be concluded that the average lifetime of a circulated \$1 bill differs from 18 months?

2. **Travel Times to Work** Based on information from the U.S. Census Bureau, the mean travel time to work in minutes for all workers 16 years old and older was 25.3 minutes. A large company with offices in several states randomly sampled 100 of its workers to ascertain their commuting times. The sample mean was 23.9 minutes, and the population standard deviation is 6.39 minutes. At the 0.01 level of significance, can it be concluded that the mean commuting time is less for this particular company?

   Source: factfinder.census.gov

3. **Debt of College Graduates** A random sample of the average debt (in dollars) at graduation from 30 of the top 100 public colleges and universities is listed below. Is there sufficient evidence at $\alpha = 0.01$ to conclude that the population mean debt at graduation is less than \$18,000? Assume $\sigma = \$2605$.

16,012	15,784	16,597	18,105	12,665	14,734
17,225	16,953	15,309	15,297	14,437	14,835
13,607	13,374	19,410	18,385	22,312	16,656
20,142	17,821	12,701	22,400	15,730	17,673
18,978	13,661	12,580	14,392	16,000	15,176

Source: www.Kiplinger.com

4. **Time Until Indigestion Relief** An advertisement claims that Fasto Stomach Calm will provide relief from indigestion in less than 10 minutes. For a test of the claim, 35 randomly selected individuals were given the product; the average time until relief was 9.25 minutes. From past studies, the standard deviation of the population is known to be 2 minutes. Can you conclude that the claim is justified? Find the *P*-value and let $\alpha = 0.05$.

5. **Shopper Purchases** A shopper in a grocery store purchases 22 items on average. A store manager believes that the mean is greater than 22. She surveyed a randomly selected group of 36 customers and found that the mean of the number of items purchased was 23.2. The population standard deviation is 3.7 items. At $\alpha = 0.05$, can it be concluded that the mean is greater than 22 items?

Section 8–3

6. **Trifecta Winnings** A random sample of $1 Trifecta tickets at a local racetrack paid the following amounts (in dollars and cents). Is there sufficient evidence to conclude that the average Trifecta winnings exceed $50? Use $\alpha = 0.10$. Assume the variable is normally distributed.

8.90	141.00	72.70	32.40
70.20	48.60	75.30	19.00
15.00	83.00	59.20	190.10
29.10			

7. **Weights of Men's Soccer Shoes** Is lighter better? A random sample of men's soccer shoes from an international catalog had the following weights (in ounces).

10.8	9.8	8.8	9.6	9.9
10	8.4	9.6	10	9.4
9.8	9.4	9.8		

At $\alpha = 0.05$, can it be concluded that the average weight is less than 10 ounces? Assume the variable is normally distributed.

8. **Whooping Crane Eggs** Once down to about 15, the world's only wild flock of whooping cranes now numbers a record 237 birds in its Texas Coastal Bend wintering ground (www.SunHerald.com). The average whooping crane egg weighs 208 grams. A new batch of randomly selected eggs was recently weighed, and their weights are listed. At $\alpha = 0.01$, is there sufficient evidence to conclude that the weight is greater than 208 grams? Assume the variable is normally distributed.

210	208.5	211.6	212	210.3
210.2	209	206.4	209.7	

Source: http://www.pwrc.usgs.gov/cranes.htm

Section 8–4

9. **Medical School Choices** It has been reported that one in four medical doctors received their degrees from foreign schools. A hospital researcher believes that the percentage is less than 25%. A random survey of 100 medical doctors found that 19 received their degrees from a foreign school. At $\alpha = 0.05$, is there enough evidence to support the researcher's claim?

10. **Federal Prison Populations** Nationally 60.2% of federal prisoners are serving time for drug offenses. A warden feels that in his prison the percentage is even higher. He surveys 400 randomly selected inmates' records and finds that 260 of the inmates are drug offenders. At $\alpha = 0.05$, is he correct?

Source: New York Times Almanac.

11. **Free School Lunches** It has been reported that 59.3% of U.S. school lunches served are free or at a reduced price. A random sample of 300 children in a large metropolitan area indicated that 156 of them received lunch free or at a reduced price. At the 0.01 level of significance, is there sufficient evidence to conclude that the proportion is less than 59.3%?

Source: www.fns.usda.gov

12. **MP3 Ownership** An MP3 manufacturer claims that 65% of teenagers 13 to 16 years old have their own MP3 players. A researcher wishes to test the claim and selects a random sample of 80 teenagers. She finds that 57 have their own MP3 players. At $\alpha = 0.05$, should the claim be rejected? Use the *P*-value method.

13. **Alcohol and Tobacco Use by High School Students** The use of both alcohol and tobacco by high school seniors has declined in the last 30 years. Alcohol use is down from 68.2 to 43.1%, and the use of cigarettes by high school seniors has decreased from 36.7 to 20.4%. A random sample of 300 high school seniors from a large region indicated that 18% had used cigarettes during the 30 days prior to the survey. At the 0.05 level of significance, does this differ from the national proportion?

Source: New York Times Almanac.

14. **Men Aged 65 and Over in the Labor Force** Of men aged 65 and over 20.5% are still in the U.S. labor force. A random sample of 120 retired male teachers indicated that 38 were still working. Use both a confidence interval and a hypothesis test. Test the claim that the proportion is greater than 20.5% at $\alpha = 0.10$.

Section 8–5

15. **Fuel Consumption** The standard deviation of the fuel consumption of a certain automobile is hypothesized to be greater than or equal to 4.3 miles per gallon. A random sample of 20 automobiles produced a standard deviation of 2.6 miles per gallon. Is the standard deviation really less than previously thought? Use $\alpha = 0.05$ and the *P*-value method. Assume the variable is normally distributed.

16. **Movie Admission Prices** The average movie admission price for a recent year was $7.18. The population variance was 3.81. A random sample of 15 theater

admission prices had a mean of \$8.02 with a standard deviation of \$2.08. At $\alpha = 0.05$, is there sufficient evidence to conclude a difference from the population variance? Assume the variable is normally distributed.

Source: New York Times Almanac.

17. **Games Played by NBA Scoring Leaders** A random sample of the number of games played by individual NBA scoring leaders is shown. Is there sufficient evidence to conclude that the variance in games played differs from 40? Use $\alpha = 0.05$. Assume the variable is normally distributed.

72	79	80	74	82
79	82	78	60	75

Source: Time Almanac.

18. **Fuel Consumption** The standard deviation of fuel consumption of a manufacturer's sport utility vehicle is hypothesized to be 3.3 miles per gallon. A random sample of 18 vehicles has a standard deviation of 2.8 miles per gallon. At $\alpha = 0.10$, is the claim believable?

Section 8–6

19. **Plant Leaf Lengths** A biologist knows that the average length of a leaf of a certain full-grown plant is 4 inches. The standard deviation of the population is 0.6 inch. A random sample of 20 leaves of that type of plant given a new type of plant food had an average length of 4.2 inches. Is there reason to believe that the new food is responsible for a change in the growth of the leaves? Use $\alpha = 0.01$. Find the 99% confidence interval of the mean. Do the results concur? Explain. Assume that the variable is approximately normally distributed.

20. **Tire Inflation** To see whether people are keeping their car tires inflated to the correct level of 35 pounds per square inch (psi), a tire company manager selects a random sample of 36 tires and checks the pressure. The mean of the sample is 33.5 psi, and the population standard deviation is 3 psi. Are the tires properly inflated? Use $\alpha = 0.10$. Find the 90% confidence interval of the mean. Do the results agree? Explain.

STATISTICS TODAY

How Much Better Is Better? —Revisited

Now that you have learned the techniques of hypothesis testing presented in this chapter, you realize that the difference between the sample mean and the population mean must be *significant* before you can conclude that the students really scored above average. The superintendent should follow the steps in the hypothesis-testing procedure and be able to reject the null hypothesis before announcing that his students scored higher than average.

Data Analysis

The Data Bank is found in Appendix B, or on the World Wide Web by following links from www.mhhe.com/math/stats/bluman/

1. From the Data Bank, select a random sample of at least 30 individuals, and test one or more of the following hypotheses by using the z test. Use $\alpha = 0.05$.
   - *a.* For serum cholesterol, H_0: $\mu = 220$ milligram percent (mg%). Use $\sigma = 5$.
   - *b.* For systolic pressure, H_0: $\mu = 120$ millimeters of mercury (mm Hg). Use $\sigma = 13$.
   - *c.* For IQ, H_0: $\mu = 100$. Use $\sigma = 15$.
   - *d.* For sodium level, H_0: $\mu = 140$ milliequivalents per liter (mEq/l). Use $\sigma = 6$.
2. Select a random sample of 15 individuals and test one or more of the hypotheses in Exercise 1 by using the t test. Use $\alpha = 0.05$.
3. Select a random sample of at least 30 individuals, and using the z test for proportions, test one or more of the following hypotheses. Use $\alpha = 0.05$.
   - *a.* For educational level, H_0: $p = 0.50$ for level 2.
   - *b.* For smoking status, H_0: $p = 0.20$ for level 1.
   - *c.* For exercise level, H_0: $p = 0.10$ for level 1.
   - *d.* For gender, H_0: $p = 0.50$ for males.
4. Select a sample of 20 individuals and test the hypothesis H_0: $\sigma^2 = 225$ for IQ level. Use $\alpha = 0.05$. Assume the variable is normally distributed.
5. Using the data from Data Set XIII, select a sample of 10 hospitals, and test H_0: $\mu = 250$ and H_1: $\mu < 250$ for the number of beds. Use $\alpha = 0.05$. Assume the variable is normally distributed.
6. Using the data obtained in Exercise 5, test the hypothesis H_0: $\sigma \geq 150$. Use $\alpha = 0.05$. Assume the variable is normally distributed.

Chapter Quiz

Determine whether each statement is true or false. If the statement is false, explain why.

1. No error is committed when the null hypothesis is rejected when it is false.

2. When you are conducting the *t* test, the population must be approximately normally distributed.

3. The test value separates the critical region from the noncritical region.

4. The values of a chi-square test cannot be negative.

5. The chi-square test for variances is always one-tailed.

Select the best answer.

6. When the value of α is increased, the probability of committing a type I error is

a. Decreased
b. Increased
c. The same
d. None of the above

7. If you wish to test the claim that the mean of the population is 100, the appropriate null hypothesis is

a. $\overline{X} = 100$
b. $\mu \geq 100$
c. $\mu \leq 100$
d. $\mu = 100$

8. The degrees of freedom for the chi-square test for variances or standard deviations are

a. 1
b. n
c. $n - 1$
d. None of the above

9. For the *t* test, one uses _______ instead of σ.

a. n
b. s
c. χ^2
d. t

Complete the following statements with the best answer.

10. Rejecting the null hypothesis when it is true is called a(n) _______ error.

11. The probability of a type II error is referred to as _______.

12. A conjecture about a population parameter is called a(n) _______.

13. To test the claim that the mean is greater than 87, you would use a(n) _______-tailed test.

14. The degrees of freedom for the *t* test are _________.

For the following exercises where applicable:

a. State the hypotheses and identify the claim.
b. Find the critical value(s).
c. Compute the test value.
d. Make the decision.
e. Summarize the results.

Use the traditional method of hypothesis testing unless otherwise specified. Assume all variables are normally distributed.

15. Ages of Professional Women A sociologist wishes to see if it is true that for a certain group of professional women, the average age at which they have their first child is 28.6 years. A random sample of 36 women is selected, and their ages at the birth of their first child are recorded. At $\alpha = 0.05$, does the evidence refute the sociologist's assertion? Assume $\sigma = 4.18$.

32	28	26	33	35	34
29	24	22	25	26	28
28	34	33	32	30	29
30	27	33	34	28	25
24	33	25	37	35	33
34	36	38	27	29	26

16. Home Closing Costs A real estate agent believes that the average closing cost of purchasing a new home is \$6500 over the purchase price. She selects 40 new home sales at random and finds that the average closing costs are \$6600. The standard deviation of the population is \$120. Test her belief at $\alpha = 0.05$.

17. Chewing Gum Use A recent study stated that if a person chewed gum, the average number of sticks of gum he or she chewed daily was 8. To test the claim, a researcher selected a random sample of 36 gum chewers and found the mean number of sticks of gum chewed per day was 9. The standard deviation of the population is 1. At $\alpha = 0.05$, is the number of sticks of gum a person chews per day actually greater than 8?

18. Hotel Rooms A travel agent claims that the average of the number of rooms in hotels in a large city is 500. At $\alpha = 0.01$, is the claim realistic? The data for a random sample of seven hotels are shown.

713	300	292	311	598	401	618

Give a reason why the claim might be deceptive.

19. Heights of Models In a New York modeling agency, a researcher wishes to see if the average height of female models is really less than 67 inches, as the chief claims. A random sample of 20 models has an average height of 65.8 inches. The standard deviation of the sample is 1.7 inches. At $\alpha = 0.05$, is the average height of the models really less than 67 inches? Use the *P*-value method.

20. **Experience of Taxi Drivers** A taxi company claims that its drivers have an average of at least 12.4 years' experience. In a study of 15 randomly selected taxi drivers, the average experience was 11.2 years. The standard deviation was 2. At $\alpha = 0.10$, is the number of years' experience of the taxi drivers really less than the taxi company claimed?

21. **Ages of Robbery Victims** A recent study in a small city stated that the average age of robbery victims was 63.5 years. A random sample of 20 recent victims had a mean of 63.7 years and a standard deviation of 1.9 years. At $\alpha = 0.05$, is the average age higher than originally believed? Use the *P*-value method.

22. **First-Time Marriages** A magazine article stated that the average age of women who are getting married for the first time is 26 years. A researcher decided to test this hypothesis at $\alpha = 0.02$. She selected a random sample of 25 women who were recently married for the first time and found the average was 25.1 years. The standard deviation was 3 years. Should the null hypothesis be rejected on the basis of the sample?

23. **Survey on Vitamin Usage** A survey in *Men's Health* magazine reported that 39% of cardiologists said that they took vitamin E supplements. To see if this is still true, a researcher randomly selected 100 cardiologists and found that 36 said that they took vitamin E supplements. At $\alpha = 0.05$, test the claim that 39% of the cardiologists took vitamin E supplements.

24. **Breakfast Survey** A dietitian read in a survey that at least 55% of adults do not eat breakfast at least 3 days a week. To verify this, she selected a random sample of 80 adults and asked them how many days a week they skipped breakfast. A total of 50% responded that they skipped breakfast at least 3 days a week. At $\alpha = 0.10$, test the claim.

25. **Caffeinated Beverage Survey** A Harris Poll found that 35% of people said that they drink a caffeinated beverage to combat midday drowsiness. A recent survey found that 19 out of 48 randomly selected people stated that they drank a caffeinated beverage to combat midday drowsiness. At $\alpha = 0.02$, is the claim of the percentage found in the Harris Poll believable?

26. **Radio Ownership** A magazine claims that 75% of all teenage boys have their own radios. A researcher wished to test the claim and selected a random sample of 60 teenage boys. She found that 54 had their own radios. At $\alpha = 0.01$, should the claim be rejected?

27. Find the *P*-value for the *z* test in Exercise 15.

28. Find the *P*-value for the *z* test in Exercise 16.

29. **Pages in Romance Novels** A copyeditor thinks the standard deviation for the number of pages in a romance novel is greater than 6. A random sample of 25 novels has a standard deviation of 9 pages. At $\alpha = 0.05$, is it higher, as the editor hypothesized?

30. **Seed Germination Times** It has been hypothesized that the standard deviation of the germination time of radish seeds is 8 days. The standard deviation of a random sample of 60 radish plants' germination times was 6 days. At $\alpha = 0.01$, test the claim.

31. **Pollution By-products** The standard deviation of the pollution by-products released in the burning of 1 gallon of gas is 2.3 ounces. A random sample of 20 automobiles tested produced a standard deviation of 1.9 ounces. Is the standard deviation really less than previously thought? Use $\alpha = 0.05$.

32. **Strength of Wrapping Cord** A manufacturer claims that the standard deviation of the strength of wrapping cord is 9 pounds. A random sample of 10 wrapping cords produced a standard deviation of 11 pounds. At $\alpha = 0.05$, test the claim. Use the *P*-value method.

33. Find the 90% confidence interval of the mean in Exercise 15. Is μ contained in the interval?

34. Find the 95% confidence interval for the mean in Exercise 16. Is μ contained in the interval?

Critical Thinking Challenges

The power of a test $(1 - \beta)$ can be calculated when a specific value of the mean is hypothesized in the alternative hypothesis; for example, let H_0: $\mu = 50$ and let H_1: $\mu = 52$. To find the power of a test, it is necessary to find the value of β. This can be done by the following steps:

Step 1 For a specific value of α find the corresponding value of $\overline{X}$, using $z = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}}$, where μ is the hypothesized value given in H_0. Use a right-tailed test.

Step 2 Using the value of $\overline{X}$ found in step 1 and the value of μ in the alternative hypothesis, find the area corresponding to z in the formula $z = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}}$.

Step 3 Subtract this area from 0.5000. This is the value of β.

Step 4 Subtract the value of β from 1. This will give you the power of a test. See Figure 8–41.

1. Find the power of a test, using the hypotheses given previously and $\alpha = 0.05$, $\sigma = 3$, and $n = 30$.
2. Select several other values for μ in H_1 and compute the power of the test. Generalize the results.

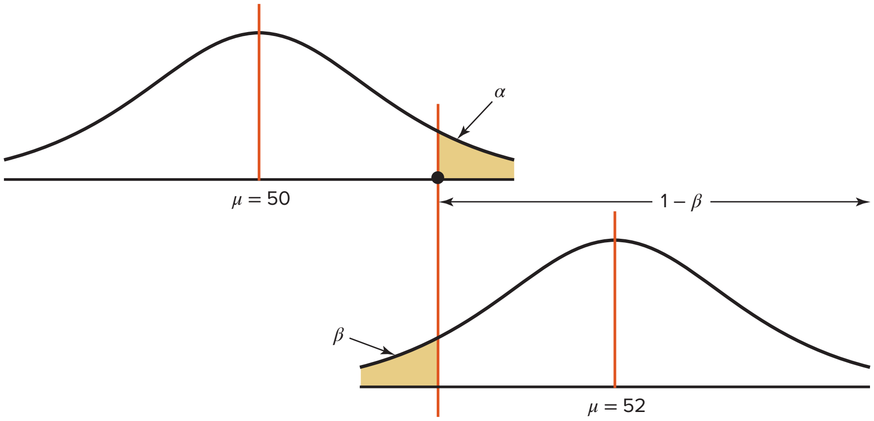


FIGURE 8–41
Relationship Among α, β, and the Power of a Test

Data Projects

Use a significance level of 0.05 for all tests below.

1. **Business and Finance** Use the Dow Jones Industrial stocks in data project 1 of Chapter 7 as your data set. Find the gain or loss for each stock over the last quarter. Test the claim that the mean is that the stocks broke even (no gain or loss indicates a mean of 0).
2. **Sports and Leisure** Use the most recent NFL season for your data. For each team, find the quarterback rating for the number one quarterback. Test the claim that the mean quarterback rating for a number one quarterback is more than 80.
3. **Technology** Use your last month's itemized cell phone bill for your data. Determine the percentage of your text messages that were outgoing. Test the claim that a majority of your text messages were outgoing. Determine the mean, median, and standard deviation for the length of a call. Test the claim that the mean length of a call is longer than the value you found for the median length.
4. **Health and Wellness** Use the data collected in data project 4 of Chapter 7 for this exercise. Test the claim that the mean body temperature is less than 98.6 degrees Fahrenheit.
5. **Politics and Economics** Use the most recent results of the Presidential primary elections for both parties. Determine what percentage of voters in your state voted for the eventual Democratic nominee for President and what percentage voted for the eventual Republican nominee. Test the claim that a majority of your state favored the candidate who won the nomination for each party.
6. **Your Class** Use the data collected in data project 6 of Chapter 7 for this exercise. Test the claim that the mean BMI for a student is more than 25.

Answers to Applying the Concepts

Section 8–1 Eggs and Your Health

1. The study was prompted by claims that linked eating eggs to high blood serum cholesterol.
2. The population under study is people in general.
3. A sample of 500 subjects was collected.
4. The hypothesis was that eating eggs did not increase blood serum cholesterol.
5. Blood serum cholesterol levels were collected.
6. Most likely, but we are not told which test.
7. The conclusion was that eating a moderate amount of eggs will not significantly increase blood serum cholesterol level.

Section 8–2 Car Thefts

1. The hypotheses are H_0: $\mu = 44$ and H_1: $\mu \neq 44$.
2. This sample can be considered large for our purposes.
3. The variable needs to be normally distributed.
4. We will use a z distribution.

5. Since we are interested in whether the car theft rate has changed, we use a two-tailed test.
6. Answers may vary. At the $\alpha = 0.05$ significance level, the critical values are $z = \pm 1.96$.
7. The sample mean is $\overline{X} = 55.97$ and the population standard deviation is 30.30. Our test statistic is $z = \dfrac{55.97 - 44}{30.30/\sqrt{36}} = 2.37$.
8. Since $2.37 > 1.96$, we reject the null hypothesis.
9. There is enough evidence to conclude that the car theft rate has changed.
10. Answers will vary. Based on our sample data, it appears that the car theft rate has changed from 44 vehicles per 10,000 people. In fact, the data indicate that the car theft rate has increased.
11. Based on our sample, we would expect 55.97 car thefts per 10,000 people, so we would expect $(55.97)(5) = 279.85$, or about 280, car thefts in the city.

Section 8–3 How Much Nicotine Is in Those Cigarettes?

1. We have $15 - 1 = 14$ degrees of freedom.
2. This is a t test.
3. We are only testing one sample.
4. This is a right-tailed test, since the hypotheses of the tobacco company are H_0: $\mu = 40$ and H_1: $\mu > 40$.
5. The P-value is 0.008, which is less than the significance level of 0.01. We reject the tobacco company's claim.
6. Since the test statistic (2.72) is greater than the critical value (2.62), we reject the tobacco company's claim.
7. There is no conflict in this output, since the results based on the P-value and the test statistic value agree.
8. Answers will vary. It appears that the company's claim is false and that there is more than 40 mg of nicotine in its cigarettes.

Section 8–4 Quitting Smoking

1. The statistical hypotheses were that StopSmoke helps more people quit smoking than the other leading brands.
2. The null hypotheses were that StopSmoke has the same effectiveness as or is not as effective as the other leading brands.
3. The alternative hypotheses were that StopSmoke helps more people quit smoking than the other leading brands. (The alternative hypotheses are the statistical hypotheses.)
4. No statistical tests were run that we know of.
5. Had tests been run, they would have been one-tailed tests.
6. Some possible significance levels are 0.01, 0.05, and 0.10.
7. A type I error would be to conclude that StopSmoke is better when it really is not.
8. A type II error would be to conclude that StopSmoke is not better when it really is.
9. These studies proved nothing. Had statistical tests been used, we could have tested the effectiveness of StopSmoke.
10. Answers will vary. One possible answer is that more than likely the statements are talking about practical significance and not statistical significance, since we have no indication that any statistical tests were conducted.

Section 8–5 Testing Gas Mileage Claims

1. The hypotheses are H_0: $\mu = 28$ and H_1: $\mu < 28$. The value of our test statistic is $t = -1.96$, and the associated P-value is 0.0287. We would reject Chrysler's claim at $\alpha = 0.05$ that the Dodge Caravans are getting 28 mpg.
2. The hypotheses are H_0: $\sigma = 2.1$ and H_1: $\sigma > 2.1$. The value of our test statistic is $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2} = \dfrac{(39)4.2^2}{2.1^2} = 156$, and the associated P-value is approximately zero. We would reject Chrysler's claim that the standard deviation is no more than 2.1 mpg.
3. Answers will vary. It is recommended that Chrysler lower its claim about the highway miles per gallon of the Dodge Caravans. Chrysler should also try to reduce variability in miles per gallon and provide confidence intervals for the highway miles per gallon.
4. Answers will vary. There are cases when a mean may be fine, but if there is a lot of variability about the mean, there will be complaints (due to the lack of consistency).

Section 8–6 Consumer Protection Agency Complaints

1. Answers will vary.
2. Answers will vary.
3. Answers will vary.
4. Answers will vary.
5. Answers will vary.
6. Answers will vary.

9

Testing the Difference Between Two Means, Two Proportions, and Two Variances

© Fuse/Corbis/Getty Images RF

STATISTICS TODAY

To Vaccinate or Not to Vaccinate? Small versus Large Nursing Homes

Influenza is a serious disease among the elderly, especially those living in nursing homes. Those residents are more susceptible to influenza than elderly persons living in the community because the former are usually older and more debilitated, and they live in a closed environment where they are exposed more so than community residents to the virus if it is introduced into the home. Three researchers decided to investigate the use of vaccine and its value in determining outbreaks of influenza in small nursing homes.

These researchers surveyed 83 randomly selected licensed homes in seven counties in Michigan. Part of the study consisted of comparing the number of people being vaccinated in small nursing homes (100 or fewer beds) with the number in larger nursing homes (more than 100 beds). Unlike the statistical methods presented in Chapter 8, these researchers used the techniques explained in this chapter to compare two sample proportions to see if there was a significant difference in the vaccination rates of patients in small nursing homes compared to those in large nursing homes. See Statistics Today—Revisited at the end of the chapter.

Source: Nancy Arden, Arnold S. Monto, and Suzanne E. Ohmit, "Vaccine Use and the Risk of Outbreaks in a Sample of Nursing Homes During an Influenza Epidemic," *American Journal of Public Health* 85, no. 3, pp. 399–401. Copyright by the American Public Health Association.

OUTLINE

Introduction

9–1 Testing the Difference Between Two Means: Using the *z* Test

9–2 Testing the Difference Between Two Means of Independent Samples: Using the *t* Test

9–3 Testing the Difference Between Two Means: Dependent Samples

9–4 Testing the Difference Between Proportions

9–5 Testing the Difference Between Two Variances

Summary

OBJECTIVES

After completing this chapter, you should be able to:

① Test the difference between two means, using the *z* test.

② Test the difference between two means for independent samples, using the *t* test.

③ Test the difference between two means for dependent samples.

④ Test the difference between two proportions.

⑤ Test the difference between two variances or standard deviations.

Introduction

The basic concepts of hypothesis testing were explained in Chapter 8. With the z, t, and χ^2 tests, a sample mean, variance, or proportion can be compared to a specific population mean, variance, or proportion to determine whether the null hypothesis should be rejected.

There are, however, many instances when researchers wish to compare two sample means, using experimental and control groups. For example, the average lifetimes of two different brands of bus tires might be compared to see whether there is any difference in tread wear. Two different brands of fertilizer might be tested to see whether one is better than the other for growing plants. Or two brands of cough syrup might be tested to see whether one brand is more effective than the other.

In the comparison of two means, the same basic steps for hypothesis testing shown in Chapter 8 are used, and the z and t tests are also used. When comparing two means by using the t test, the researcher must decide if the two samples are *independent* or *dependent*. The concepts of independent and dependent samples will be explained in Sections 9–2 and 9–3.

The z test can be used to compare two proportions, as shown in Section 9–4. Finally, two variances can be compared by using an F test as shown in Section 9–5.

9–1 Testing the Difference Between Two Means: Using the z Test

OBJECTIVE ①

Test the difference between two means, using the z test.

Suppose a researcher wishes to determine whether there is a difference in the average age of nursing students who enroll in a nursing program at a community college and those who enroll in a nursing program at a university. In this case, the researcher is not interested in the average age of all beginning nursing students; instead, he is interested in *comparing* the means of the two groups. His research question is, Does the mean age of nursing students who enroll at a community college differ from the mean age of nursing students who enroll at a university? Here, the hypotheses are

$$H_0\colon \mu_1 = \mu_2$$
$$H_1\colon \mu_1 \neq \mu_2$$

where

μ_1 = mean age of all beginning nursing students at a community college
μ_2 = mean age of all beginning nursing students at a university

Another way of stating the hypotheses for this situation is

$$H_0\colon \mu_1 - \mu_2 = 0$$
$$H_1\colon \mu_1 - \mu_2 \neq 0$$

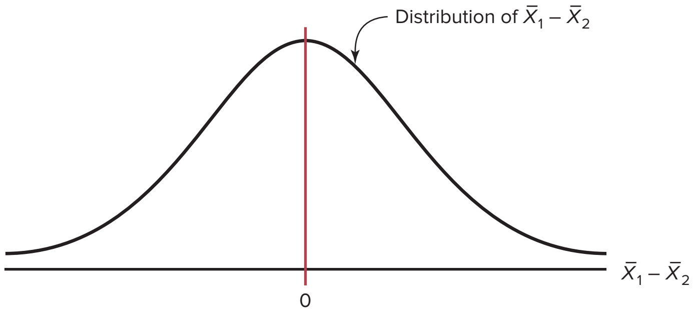
If there is no difference in population means, subtracting them will give a difference of zero. If they are different, subtracting will give a number other than zero. Both methods of stating hypotheses are correct; however, the first method will be used in this text.

If two samples are *independent* of each other, the subjects selected for the first sample in no way influence the way the subjects are selected in the second sample. For example, if a group of 50 people were randomly divided into two groups of 25 people each in order to test the effectiveness of a new drug, where one group gets the drug and the other group gets a placebo, the samples would be independent of each other.

On the other hand, two samples would be *dependent* if the selection of subjects for the first group in some way influenced the selection of subjects for the other group. For example, suppose you wanted to determine if a person's right foot was slightly larger than his or her left foot. In this case, the samples are dependent because once you selected a

FIGURE 9–1

Differences of Means of Pairs of Samples



person's right foot for sample 1, you must select his or her left foot for sample 2 because you are using the same person for both feet.

Before you can use the *z* test to test the difference between two independent sample means, you must make sure that the following assumptions are met.

Assumptions for the *z* Test to Determine the Difference Between Two Means

1. Both samples are random samples.
2. The samples must be independent of each other. That is, there can be no relationship between the subjects in each sample.
3. The standard deviations of both populations must be known; and if the sample sizes are less than 30, the populations must be normally or approximately normally distributed.

In this book, the assumptions will be stated in the exercises; however, when encountering statistics in other situations, you must check to see that these assumptions have been met before proceeding.

The theory behind testing the difference between two means is based on selecting pairs of samples and comparing the means of the pairs. The population means need not be known.

All possible pairs of samples are taken from populations. The means for each pair of samples are computed and then subtracted, and the differences are plotted. If both populations have the same mean, then most of the differences will be zero or close to zero. Occasionally, there will be a few large differences due to chance alone, some positive and others negative. If the differences are plotted, the curve will be shaped like a normal distribution and have a mean of zero, as shown in Figure 9–1.

Unusual Stats

Adult children who live with their parents spend more than 2 hours a day doing household chores. According to a study, daughters contribute about 17 hours a week and sons about 14.4 hours.

The variance of the difference $\overline{X}_1 - \overline{X}_2$ is equal to the sum of the individual variances of $\overline{X}_1$ and $\overline{X}_2$. That is,

$$\sigma^2_{\overline{X}_1 - \overline{X}_2} = \sigma^2_{\overline{X}_1} + \sigma^2_{\overline{X}_2}$$

where $\sigma^2_{\overline{X}_1} = \dfrac{\sigma_1^2}{n_1}$ and $\sigma^2_{\overline{X}_2} = \dfrac{\sigma_2^2}{n_2}$

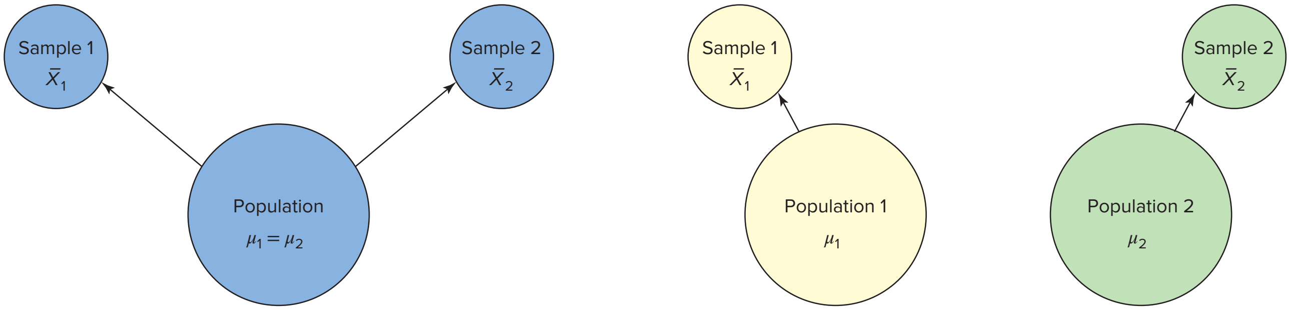
So the standard deviation of $\overline{X}_1 - \overline{X}_2$ is

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Formula for the *z* Test for Comparing Two Means from Independent Populations

$$z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

FIGURE 9–2 Hypothesis-Testing Situations in the Comparison of Means



(a) Difference is not significant. The means of the populations are the same. Do not reject H_0: $\mu_1 = \mu_2$ since $\overline{X}_1 - \overline{X}_2$ is not significant.

(b) Difference is significant. The means of the populations are different. Reject H_0: $\mu_1 = \mu_2$ since $\overline{X}_1 - \overline{X}_2$ is significant.

This formula is based on the general format of

$$\text{Test value} = \frac{(\text{observed value}) - (\text{expected value})}{\text{standard error}}$$

where $\overline{X}_1 - \overline{X}_2$ is the observed difference, and the expected difference $\mu_1 - \mu_2$ is zero when the null hypothesis is $\mu_1 = \mu_2$, since that is equivalent to $\mu_1 - \mu_2 = 0$. Finally, the standard error of the difference is

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

In the comparison of two sample means, the difference may be due to chance, in which case the null hypothesis will not be rejected and the researcher can assume that the means of the populations are basically the same. The difference in this case is not significant. See Figure 9–2(a). On the other hand, if the difference is significant, the null hypothesis is rejected and the researcher can conclude that the population means are different. See Figure 9–2(b).

These tests can also be one-tailed, using the following hypotheses:

Right-tailed			Left-tailed		
H_0: $\mu_1 = \mu_2$ H_1: $\mu_1 > \mu_2$	or	H_0: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 > 0$	H_0: $\mu_1 = \mu_2$ H_1: $\mu_1 < \mu_2$	or	H_0: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 < 0$

The same critical values used in Section 8–2 are used here. They can be obtained from Table E in Appendix A.

The basic format for hypothesis testing using the traditional method is reviewed here.

Step 1 State the hypotheses and identify the claim.
Step 2 Find the critical value(s).
Step 3 Compute the test value.
Step 4 Make the decision.
Step 5 Summarize the results.

EXAMPLE 9–1 Leisure Time

A study using two random samples of 35 people each found that the average amount of time those in the age group of 26–35 years spent per week on leisure activities was 39.6 hours, and those in the age group of 46–55 years spent 35.4 hours. Assume that the population standard deviation for those in the first age group found by previous studies is 6.3 hours, and the population standard deviation of those in the second group found by previous studies was 5.8 hours. At $\alpha = 0.05$, can it be concluded that there is a significant difference in the average times each group spends on leisure activities?

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon \mu_1 = \mu_2 \qquad \text{and} \qquad H_1\colon \mu_1 \neq \mu_2 \text{ (claim)}$$

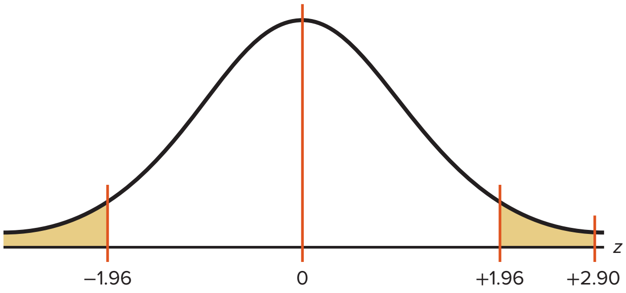
Step 2 Find the critical values. Since $\alpha = 0.05$, the critical values are $+1.96$ and -1.96.

Step 3 Compute the test value.

$$z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} = \frac{(39.6 - 35.4) - 0}{\sqrt{\dfrac{6.3^2}{35} + \dfrac{5.8^2}{35}}} = \frac{4.2}{1.447} = 2.90$$

Step 4 Make the decision. Reject the null hypothesis at $\alpha = 0.05$ since $2.90 > 1.96$. See Figure 9–3.

FIGURE 9–3 Critical and Test Values for Example 9–1



Step 5 Summarize the results. There is enough evidence to support the claim that the means are not equal. That is, the average of the times spent on leisure activities is different for the groups.

The P-values for this test can be determined by using the same procedure shown in Section 8–2. For example, if the test value for a two-tailed test is 2.90, then the P-value obtained from Table E is 0.0038. This value is obtained by looking up the area for $z = 2.90$, which is 0.9981. Then 0.9981 is subtracted from 1.0000 to get 0.0019. Finally, this value is doubled to get 0.0038 since the test is two-tailed. If $\alpha = 0.05$, the decision would be to reject the null hypothesis, since $P\text{-value} < \alpha$ (that is, $0.0038 < 0.05$). *Note:* The P-value obtained on the TI-84 is 0.0037.

The P-value method for hypothesis testing for this chapter also follows the same format as stated in Chapter 8. The steps are reviewed here.

Step 1 State the hypotheses and identify the claim.
Step 2 Compute the test value.
Step 3 Find the P-value.
Step 4 Make the decision.
Step 5 Summarize the results.

Example 9–2 illustrates these steps.

EXAMPLE 9–2 College Sports Offerings

A researcher hypothesizes that the average number of sports that colleges offer for males is greater than the average number of sports that colleges offer for females. A random sample of the number of sports offered by colleges for males and females is shown. At $\alpha = 0.10$, is there enough evidence to support the claim? Assume σ_1 and $\sigma_2 = 3.3$.

Males					Females				
6	11	11	8	15	6	8	11	13	8
6	14	8	12	18	7	5	13	14	6
6	9	5	6	9	6	5	5	7	6
6	9	18	7	6	10	7	6	5	5
15	6	11	5	5	16	10	7	8	5
9	9	5	5	8	7	5	5	6	5
8	9	6	11	6	9	18	13	7	10
9	5	11	5	8	7	8	5	7	6
7	7	5	10	7	11	4	6	8	7
10	7	10	8	11	14	12	5	8	5

Source: USA TODAY.

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon \mu_1 = \mu_2 \qquad \text{and} \qquad H_1\colon \mu_1 > \mu_2 \text{ (claim)}$$

Step 2 Compute the test value. Using a calculator or the formula in Chapter 3, find the mean for each data set.

For the males $\overline{X}_1 = 8.6$ and $\sigma_1 = 3.3$

For the females $\overline{X}_2 = 7.9$ and $\sigma_2 = 3.3$

Substitute in the formula.

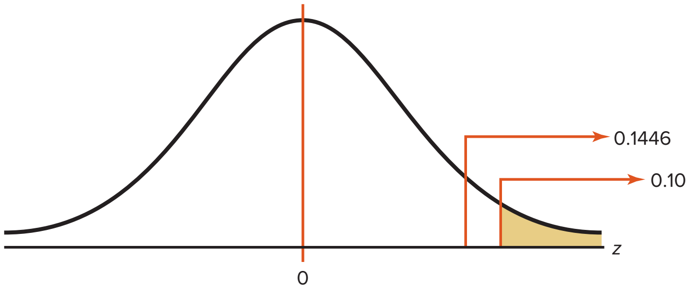
$$z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} = \frac{(8.6 - 7.9) - 0}{\sqrt{\dfrac{3.3^2}{50} + \dfrac{3.3^2}{50}}} = 1.06^*$$

Step 3 Find the *P*-value from Table E. For $z = 1.06$, the area is 0.8554, and $1.0000 - 0.8554 = 0.1446$, or a *P*-value of 0.1446.

Step 4 Make the decision. Since the *P*-value is larger than α (that is, $0.1446 > 0.10$), the decision is to not reject the null hypothesis. See Figure 9–4.

Step 5 Summarize the results. There is not enough evidence to support the claim that colleges offer more sports for males than they do for females at the 0.10 level of significance.

FIGURE 9–4 *P*-Value and α Value for Example 9–2



**Note:* Calculator results may differ due to rounding.

Sometimes, the researcher is interested in testing a specific difference in means other than zero. For example, he or she might hypothesize that the nursing students at a community college are, on average, 3.2 years older than those at a university. In this case, the hypotheses are

$$H_0\colon \mu_1 - \mu_2 = 3.2 \qquad \text{and} \qquad H_1\colon \mu_1 - \mu_2 > 3.2$$

The formula for the z test is still

$$z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

where $\mu_1 - \mu_2$ is the hypothesized difference or expected value. In this case, $\mu_1 - \mu_2 = 3.2$.

Confidence intervals for the difference between two means can also be found. When you are hypothesizing a difference of zero, if the confidence interval contains zero, the null hypothesis is not rejected. If the confidence interval does not contain zero, the null hypothesis is rejected.

Confidence intervals for the difference between two means can be found by using this formula:

Formula for the z Confidence Interval for Difference Between Two Means

$$(\overline{X}_1 - \overline{X}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\overline{X}_1 - \overline{X}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

EXAMPLE 9–3 Leisure Time

Find the 95% confidence interval for the difference between the means in Example 9–1.

SOLUTION

Substitute in the formula, using $z_{\alpha/2} = 1.96$.

$$(\overline{X}_1 - \overline{X}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\overline{X}_1 - \overline{X}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$(39.6 - 35.4) - 1.96\sqrt{\frac{6.3^2}{35} + \frac{5.8^2}{35}} < \mu_1 - \mu_2 < (39.6 - 35.4) + 1.96\sqrt{\frac{6.3^2}{35} + \frac{5.8^2}{35}}$$

$$4.2 - 2.8 < \mu_1 - \mu_2 < 4.2 + 2.8$$

$$1.4 < \mu_1 - \mu_2 < 7.0$$

(The confidence interval obtained from the TI-84 is $1.363 < \mu_1 - \mu_2 < 7.037$.)

Since the confidence interval does not contain zero, the decision is to reject the null hypothesis, which agrees with the previous result.

Applying the Concepts 9–1

Home Runs

For a sports radio talk show, you are asked to research the question whether more home runs are hit by players in the National League or by players in the American League. You decide to use the home run leaders from each league for a 40-year period as your data. The numbers are shown.

National League									
47	49	73	50	65	70	49	47	40	43
46	35	38	40	47	39	49	37	37	36
40	37	31	48	48	45	52	38	38	36
44	40	48	45	45	36	39	44	52	47

American League									
47	57	52	47	48	56	56	52	50	40
46	43	44	51	36	42	49	49	40	43
39	39	22	41	45	46	39	32	36	32
32	32	37	33	44	49	44	44	49	32

Using the data given, answer the following questions.

1. Define a population.
2. What kind of sample was used?
3. Do you feel that the samples are representative?
4. What are your hypotheses?
5. What significance level will you use?
6. What statistical test will you use?
7. What are the test results? (Assume $\sigma_1 = 8.8$ and $\sigma_2 = 7.8$.)
8. What is your decision?
9. What can you conclude?
10. Do you feel that using the data given really answers the original question asked?
11. What other data might be used to answer the question?

See page 544 for the answers.

Exercises 9–1

1. Explain the difference between testing a single mean and testing the difference between two means.

2. When a researcher selects all possible pairs of samples from a population in order to find the difference between the means of each pair, what will be the shape of the distribution of the differences when the original distributions are normally distributed? What will be the mean of the distribution? What will be the standard deviation of the distribution?

3. What three assumptions must be met when you are using the z test to test differences between two means when σ_1 and σ_2 are known?

4. Show two different ways to state that the means of two populations are equal.

For Exercises 5 through 16, perform each of the following steps.

a. State the hypotheses and identify the claim.
b. Find the critical value(s).
c. Compute the test value.
d. Make the decision.
e. Summarize the results.

Use the traditional method of hypothesis testing unless otherwise specified.

5. Recreational Time A researcher wishes to see if there is a difference between the mean number of hours per week that a family with no children participates in recreational activities and a family with children participates in recreational activities. She selects two random samples and the data are shown. At $\alpha = 0.10$, is there a difference between the means?

	$\bar{X}$	σ	n
No children	8.6	2.1	36
Children	10.6	2.7	36

6. Teachers' Salaries Teachers' Salaries New York and Massachusetts lead the list of average teacher's salaries.

The New York average is \$76,409 while teachers in Massachusetts make an average annual salary of \$73,195. Random samples of 45 teachers from each state yielded the following.

	Massachusetts	New York
Sample means	\$73,195	\$76,409
Population standard deviation	8,200	7,800

At $\alpha = 0.10$, is there a difference in means of the salaries?

Source: World Almanac.

7. **Commuting Times** The U.S. Census Bureau reports that the average commuting time for citizens of both Baltimore, Maryland, and Miami, Florida, is approximately 29 minutes. To see if their commuting times appear to be any different in the winter, random samples of 40 drivers were surveyed in each city and the average commuting time for the month of January was calculated for both cities. The results are shown. At the 0.05 level of significance, can it be concluded that the commuting times are different in the winter?

	Miami	Baltimore
Sample size	40	40
Sample mean	28.5 min	35.2 min
Population standard deviation	7.2 min	9.1 min

Source: www.census.gov

8. **Heights of 9-Year-Olds** At age 9 the average weight (21.3 kg) and the average height (124.5 cm) for both boys and girls are exactly the same. A random sample of 9-year-olds yielded these results. At $\alpha = 0.05$, do the data support the given claim that there is a difference in heights?

	Boys	Girls
Sample size	60	50
Mean height, cm	123.5	126.2
Population variance	98	120

Source: www.healthepic.com

9. **Length of Hospital Stays** The average length of "short hospital stays" for men is slightly longer than that for women, 5.2 days versus 4.5 days. A random sample of recent hospital stays for both men and women revealed the following. At $\alpha = 0.01$, is there sufficient evidence to conclude that the average hospital stay for men is longer than the average hospital stay for women?

	Men	Women
Sample size	32	30
Sample mean	5.5 days	4.2 days
Population standard deviation	1.2 days	1.5 days

Source: www.cdc.gov/nchs

10. **Home Prices** A real estate agent compares the selling prices of randomly selected homes in two municipalities in southwestern Pennsylvania to see if there is a difference. The results of the study are shown. Is there enough evidence to reject the claim that the average cost of a home in both locations is the same? Use $\alpha = 0.01$.

Scott	Ligonier
$\bar{X}_1 = \$93{,}430$*	$\bar{X}_2 = \$98{,}043$*
$\sigma_1 = \$5602$	$\sigma_2 = \$4731$
$n_1 = 35$	$n_2 = 40$

*Based on information from RealSTATs.

11. **Manual Dexterity Differences** A researcher wishes to see if there is a difference in the manual dexterity of athletes and that of band members. Two random samples of 30 are selected from each group and are given a manual dexterity test. The mean of the athletes' test was 87, and the mean of the band members' test was 92. The population standard deviation for the test is 7.2. At $\alpha = 0.01$, is there a significant difference in the mean scores?

12. **ACT Scores** A random survey of 1000 students nationwide showed a mean ACT score of 21.4. Ohio was not used. A survey of 500 randomly selected Ohio scores showed a mean of 20.8. If the population standard deviation is 3, can we conclude that Ohio is below the national average? Use $\alpha = 0.05$.

Source: Report of WFIN radio.

13. **Per Capita Income** The average per capita income for Wisconsin is reported to be \$37,314, and for South Dakota it is \$37,375—almost the same thing. A random sample of 50 workers from each state indicated the following sample statistics.

	Wisconsin	South Dakota
Size	50	50
Mean	\$40,275	\$38,750
Population standard deviation	\$10,500	\$12,500

At $\alpha = 0.05$, can we conclude a difference in means of the personal incomes?

Source: New York Times Almanac.

14. **Monthly Social Security Benefits** The average monthly Social Security benefit for a specific year for retired workers was \$954.90 and for disabled workers was \$894.10. Researchers used data from the Social Security records to test the claim that the difference in monthly benefits between the two groups was greater than \$30. Based on the following information, can the researchers' claim be supported at the 0.05 level of significance?

	Retired	Disabled
Sample size	60	60
Mean benefit	\$960.50	\$902.89
Population standard deviation	\$98	\$101

Source: New York Times Almanac.

15. **Self-Esteem Scores** In a study of a group of women science majors who remained in their profession and a group who left their profession within a few months of graduation, the researchers collected the data shown here on a self-esteem questionnaire. At $\alpha = 0.05$, can it be concluded that there is a difference in the self-esteem scores of the two groups? Use the *P*-value method.

Leavers	**Stayers**
$\overline{X}_1 = 3.05$	$\overline{X}_2 = 2.96$
$\sigma_1 = 0.75$	$\sigma_2 = 0.75$
$n_1 = 103$	$n_2 = 225$

Source: Paula Rayman and Belle Brett, "Women Science Majors: What Makes a Difference in Persistence after Graduation?" *The Journal of Higher Education.*

16. **Ages of College Students** The dean of students wants to see whether there is a significant difference in ages of resident students and commuting students. She selects a random sample of 50 students from each group. The ages are shown here. At $\alpha = 0.05$, decide if there is enough evidence to reject the claim of no difference in the ages of the two groups. Use the *P*-value method. Assume $\sigma_1 = 3.68$ and $\sigma_2 = 4.7$.

Resident students

22	25	27	23	26	28	26	24
25	20	26	24	27	26	18	19
18	30	26	18	18	19	32	23
19	19	18	29	19	22	18	22
26	19	19	21	23	18	20	18
22	21	19	21	21	22	18	20
19	23						

Commuter students

18	20	19	18	22	25	24	35
23	18	23	22	28	25	20	24
26	30	22	22	22	21	18	20
19	26	35	19	19	18	19	32
29	23	21	19	36	27	27	20
20	21	18	19	23	20	19	19
20	25						

17. **Working Breath Rate** Two random samples of 32 individuals were selected. One sample participated in an activity which simulates hard work. The average breath rate of these individuals was 21 breaths per minute. The other sample did some normal walking. The mean breath rate of these individuals was 14. Find the 90% confidence interval of the difference in the breath rates if the population standard deviation was 4.2 for breath rate per minute.

18. **Traveling Distances** Find the 95% confidence interval of the difference in the distance that day students travel to school and the distance evening students travel to school. Two random samples of 40 students are taken, and the data are shown. Find the 95% confidence interval of the difference in the means.

	$\overline{X}$	σ	n
Day students	4.7	1.5	40
Evening Students	6.2	1.7	40

19. **Literacy Scores** Adults aged 16 or older were assessed in three types of literacy: prose, document, and quantitative. The scores in document literacy were the same for 19- to 24-year-olds and for 40- to 49-year-olds. A random sample of scores from a later year showed the following statistics.

Age group	**Mean score**	**Population standard deviation**	**Sample size**
19–24	280	56.2	40
40–49	315	52.1	35

Construct a 95% confidence interval for the true difference in mean scores for these two groups. What does your interval say about the claim that there is no difference in mean scores?

Source: www.nces.ed.gov

20. **Age Differences** In a large hospital, a nursing director selected a random sample of 30 registered nurses and found that the mean of their ages was 30.2. The population standard deviation for the ages is 5.6. She selected a random sample of 40 nursing assistants and found the mean of their ages was 31.7. The population standard deviation of the ages for the assistants is 4.3. Find the 99% confidence interval of the differences in the ages.

21. **Television Watching** The average number of hours of television watched per week by women over age 55 is 48 hours. Men over age 55 watch an average of 43 hours of television per week. Random samples of 40 men and 40 women from a large retirement community yielded the following results. At the 0.01 level of significance, can it be concluded that women watch more television per week than men?

	Sample size	**Mean**	**Population standard deviation**
Women	40	48.2	5.6
Men	40	44.3	4.5

Source: World Almanac 2012.

22. **Commuting Times for College Students** The mean travel time to work for Americans is 25.3 minutes. An employment agency wanted to test the mean commuting times for college graduates and those with only some college. Thirty-five college graduates spent a mean time of 40.5 minutes commuting to work with a population variance of 67.24. Thirty workers who had completed some college had a mean commuting time of 34.8 minutes with a population variance of 39.69. At the 0.05 level of significance, can a difference in means be concluded?

Source: World Almanac 2012.

23. Store Sales A company owned two small Bath and Body Goods stores in different cities. It was desired to see if there was a difference in their mean daily sales. The following results were obtained from a random sample of daily sales over a six-week period. At $\alpha = 0.01$, can a difference in sales be concluded? Use the *P*-value method.

Store	Mean	Population standard deviation	Sample size
A	\$995	\$120	30
B	1120	250	30

24. Home Prices According to the almanac, the average sales price of a single-family home in the metropolitan Dallas/Ft. Worth/Irving, Texas, area is \$215,200. The average home price in Orlando, Florida, is \$198,000. The mean of a random sample of 45 homes in the Texas metroplex was \$216,000 with a population standard deviation of \$30,000. In the Orlando, Florida, area a sample of 40 homes had a mean price of \$203,000 with a population standard deviation of \$32,500. At the 0.05 level of significance, can it be concluded that the mean price in Dallas exceeds the mean price in Orlando? Use the *P*-value method.

Source: World Almanac.

Extending the Concepts

25. Exam Scores at Private and Public Schools A researcher claims that students in a private school have exam scores that are at most 8 points higher than those of students in public schools. Random samples of 60 students from each type of school are selected and given an exam. The results are shown. At $\alpha = 0.05$, test the claim.

Private school	Public school
$\overline{X}_1 = 110$	$\overline{X}_2 = 104$
$\sigma_1 = 15$	$\sigma_2 = 15$
$n_1 = 60$	$n_2 = 60$

26. Sale Prices for Houses The average sales price of new one-family houses in the Midwest is \$250,000 and in the South is \$253,400. A random sample of 40 houses in each region was examined with the following results. At the 0.05 level of significance, can it be concluded that the difference in mean sales price for the two regions is greater than \$3400?

	South	Midwest
Sample size	40	40
Sample mean	\$261,500	\$248,200
Population standard deviation	\$10,500	\$12,000

Source: New York Times Almanac.

27. Average Earnings for College Graduates The average earnings of year-round full-time workers with bachelor's degrees or more is \$88,641 for men and \$58,000 for women—a difference of slightly over \$30,000 a year. One hundred of each were randomly sampled, resulting in a sample mean of \$90,200 for men, and the population standard deviation is \$15,000; and a mean of \$57,800 for women, and the population standard deviation is \$12,800. At the 0.01 level of significance, can it be concluded that the difference in means is not \$30,000?

Source: New York Times Almanac.

Technology Step by Step

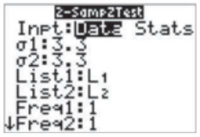
TI-84 Plus
Step by Step

Hypothesis Test for the Difference Between Two Means and *z* Distribution (Data)

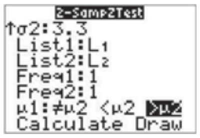
This refers to Example 9–2 in the text.

Example TI9–1

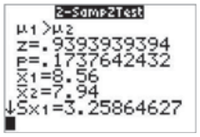
1. Enter the data values into L_1 and L_2.
2. Press **STAT** and move the cursor to TESTS.
3. Press **3** for 2-SampZTest.
4. Move the cursor to Data and press **ENTER.**
5. Type in the appropriate values.
6. Move the cursor to the appropriate alternative hypothesis and press **ENTER.**
7. Move the cursor to Calculate and press **ENTER.**



```
2-SampZTest
Inpt:Data Stats
σ1:3.3
σ2:3.3
List1:L1
List2:L2
Freq1:1
↓Freq2:1
```



```
2-SampZTest
↑σ2:3.3
List1:L1
List2:L2
Freq1:1
Freq2:1
μ1:≠μ2 <μ2 >μ2
Calculate Draw
```



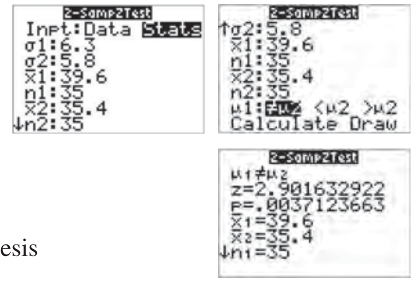
```
2-SampZTest
μ1>μ2
z=.9393939394
p=.1737642432
x̄1=8.56
x̄2=7.94
↓Sx1=3.25864627
```

Hypothesis Test for the Difference Between Two Means and *z* Distribution (Statistics)

Example TI9–2

1. Press **STAT** and move the cursor to TESTS.
2. Press **3** for 2-SampZTest.
3. Move the cursor to Stats and press **ENTER.**
4. Type in the appropriate values.
5. Move the cursor to the appropriate alternative hypothesis and press **ENTER.**
6. Move the cursor to Calculate and press **ENTER.**

This refers to Example 9–1 in the text.



```
2-SampZTest
Inpt:Data Stats
σ1:6.3
σ2:5.8
x̄1:39.6
n1:35
x̄2:35.4
↓n2:35
```

```
2-SampZTest
↑σ2:5.8
x̄1:39.6
n1:35
x̄2:35.4
n2:35
μ1:≠μ2 <μ2 >μ2
Calculate Draw
```

```
2-SampZTest
μ1≠μ2
z=2.901632922
p=.0037123663
x̄1=39.6
x̄2=35.4
↓n1=35
```

Confidence Interval for the Difference Between Two Means and *z* Distribution (Data)

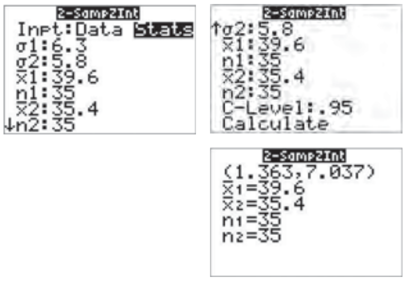
1. Enter the data values into L_1 and L_2.
2. Press **STAT** and move the cursor to TESTS.
3. Press **9** for 2-SampZInt.
4. Move the cursor to Data and press **ENTER.**
5. Type in the appropriate values.
6. Move the cursor to Calculate and press **ENTER.**

Confidence Interval for the Difference Between Two Means and *z* Distribution (Statistics)

Example TI9–3

1. Press **STAT** and move the cursor to TESTS.
2. Press **9** for 2-SampZInt.
3. Move the cursor to Stats and press **ENTER.**
4. Type in the appropriate values.
5. Move the cursor to Calculate and press **ENTER.**

This refers to Example 9–3 in the text.



```
2-SampZInt
Inpt:Data Stats
σ1:6.3
σ2:5.8
x̄1:39.6
n1:35
x̄2:35.4
↓n2:35
```

```
2-SampZInt
↑σ2:5.8
x̄1:39.6
n1:35
x̄2:35.4
n2:35
C-Level:.95
Calculate
```

```
2-SampZInt
(1.363,7.037)
x̄1=39.6
x̄2=35.4
n1=35
n2=35
```

EXCEL
Step by Step

z Test for the Difference Between Two Means

Excel has a two-sample *z* test included in the Data Analysis Add-in. To perform a *z* test for the difference between the means of two populations, given two independent samples, do this:

1. Enter the first sample data set into column A.
2. Enter the second sample data set into column B.
3. Select the Data tab from the toolbar. Then select Data Analysis.
4. In the Analysis Tools box, select *z* test: Two Sample for Means.
5. Type the ranges for the data in columns A and B and type a value (usually 0) for the Hypothesized Mean Difference.
6. Type the known population variances in for Variable 1 Variance (known) and Variable 2 Variance (known).
7. Specify the confidence level Alpha.
8. Specify a location for the output, and click [OK].

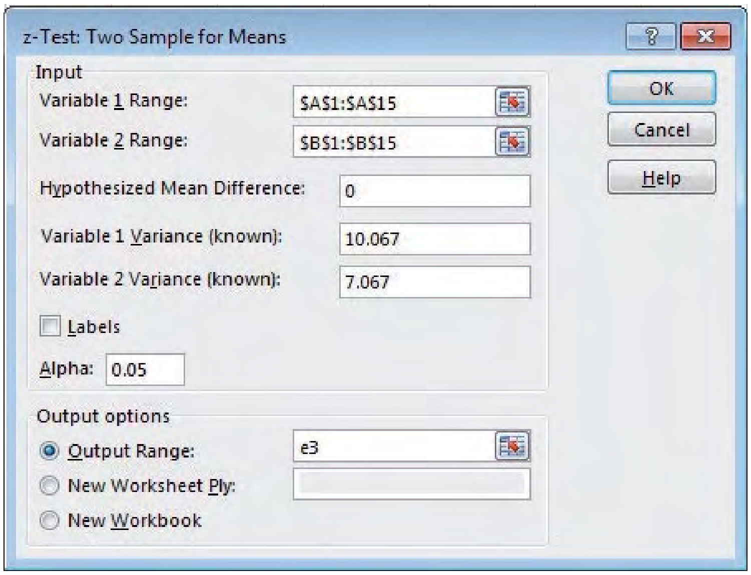
Example XL9–1

Test the claim that the two population means are equal, using the sample data provided here, at $\alpha = 0.05$. Assume the population variances are $\sigma_A^2 = 10.067$ and $\sigma_B^2 = 7.067$.

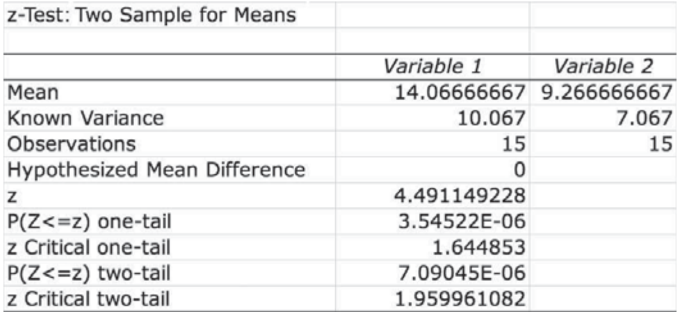
Set A	10	2	15	18	13	15	16	14	18	12	15	15	14	18	16
Set B	5	8	10	9	9	11	12	16	8	8	9	10	11	7	6

The two-sample *z* test dialog box is shown (before the variances are entered); the results appear in the table that Excel generates. Note that the *P*-value and critical *z* value are provided for both the one-tailed test and the two-tailed test. The *P*-values here are expressed in scientific notation: 7.09045E-06 = 7.09045 × 10^{-6} = 0.00000709045. Because this value is less than 0.05, we reject the null hypothesis and conclude that the population means are not equal.

Two-Sample *z* Test Dialog Box



Variable 1 Range: $A$1:$A$15
Variable 2 Range: $B$1:$B$15
Hypothesized Mean Difference: 0
Variable 1 Variance (known): 10.067
Variable 2 Variance (known): 7.067
☐ Labels
Alpha: 0.05
● Output Range: e3
○ New Worksheet Ply:
○ New Workbook



z-Test: Two Sample for Means		
	Variable 1	*Variable 2*
Mean	14.06666667	9.266666667
Known Variance	10.067	7.067
Observations	15	15
Hypothesized Mean Difference	0	
z	4.491149228	
P(Z<=z) one-tail	3.54522E-06	
z Critical one-tail	1.644853	
P(Z<=z) two-tail	7.09045E-06	
z Critical two-tail	1.959961082	

9–2 Testing the Difference Between Two Means of Independent Samples: Using the *t* Test

OBJECTIVE 2

Test the difference between two means for independent samples, using the *t* test.

In Section 9–1, the *z* test was used to test the difference between two means when the population standard deviations were known and the variables were normally or approximately normally distributed, or when both sample sizes were greater than or equal to 30. In many situations, however, these conditions cannot be met—that is, the population standard deviations are not known. In these cases, a *t* test is used to test the difference between means when the two samples are independent and when the samples are taken from two normally or approximately normally distributed populations. Samples are **independent samples** when they are not related. Also it will be assumed that the variances are not equal.

Formula for the *t* Test for Testing the Difference Between Two Means, Independent Samples

Variances are assumed to be unequal:

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

where the degrees of freedom are equal to the smaller of $n_1 - 1$ or $n_2 - 1$.

The formula

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

follows the format of

$$\text{Test value} = \frac{(\text{observed value}) - (\text{expected value})}{\text{standard error}}$$

where $\overline{X}_1 - \overline{X}_2$ is the observed difference between sample means and where the expected value $\mu_1 - \mu_2$ is equal to zero when no difference between population means is hypothesized. The denominator $\sqrt{s_1^2/n_1 + s_2^2/n_2}$ is the standard error of the difference between two means. This formula is similar to the one used when σ_1 and σ_2 are known; but when we use this t test, σ_1 and σ_2 are unknown, so s_1 and s_2 are used in the formula in place of σ_1 and σ_2. Since mathematical derivation of the standard error is somewhat complicated, it will be omitted here.

Before you can use the testing methods to determine whether two independent sample means differ when σ_1 and σ_2 are unknown, the following assumptions must be met.

Assumptions for the t Test for Two Independent Means When σ_1 and σ_2 Are Unknown

1. The samples are random samples.
2. The sample data are independent of one another.
3. When the sample sizes are less than 30, the populations must be normally or approximately normally distributed.

In this book, the assumptions will be stated in the exercises; however, when encountering statistics in other situations, you must check to see that these assumptions have been met before proceeding.

Again the hypothesis test here follows the same steps as those in Section 9–1; however, the formula uses s_1 and s_2 and Table F to get the critical values.

EXAMPLE 9–4 Work Absences

A study was done to see if there is a difference between the number of sick days men take and the number of sick days women take. A random sample of 9 men found that the mean of the number of sick days taken was 5.5. The standard deviation of the sample was 1.23. A random sample of 7 women found that the mean was 4.3 days and a standard deviation of 1.19 days. At $\alpha = 0.05$, can it be concluded that there is a difference in the means?

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon \mu_1 = \mu_2 \qquad \text{and} \qquad H_1\colon \mu_1 \neq \mu_2 \text{ (claim)}$$

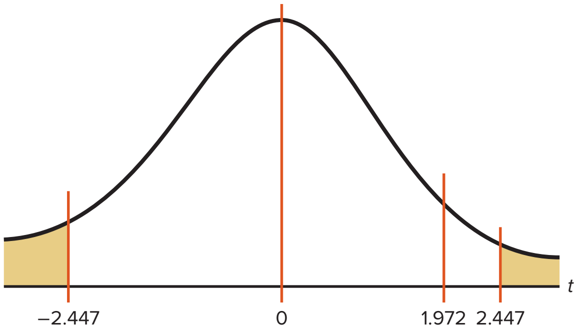
Step 2 Find the critical values. Since the test is two-tailed and $\alpha = 0.05$, the degrees of freedom are the smaller of $n_1 - 1$ and $n_2 - 1$. In this case, $n_1 - 1 = 9 - 1 = 8$ and $n_2 - 1 = 7 - 1 = 6$. So d.f. = 6. From Table F, the critical values are +2.447 and −2.447.

Step 3 Compute the test value.

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} = \frac{(5.5 - 4.3) - 0}{\sqrt{\dfrac{1.23^2}{9} + \dfrac{1.19^2}{7}}} = 1.972$$

Step 4 Make the decision. Do not reject the null hypothesis since $1.972 < 2.447$. See Figure 9–5.

FIGURE 9–5 Critical and Test Values for Example 9–4



Step 5 Summarize the results. There is not enough evidence to support the claim that the means are different.

When raw data are given in the exercises, use your calculator or the formulas in Chapter 3 to find the means and variances for the data sets. Then follow the procedures shown in this section to test the hypotheses.

Confidence intervals can also be found for the difference of two means with this formula:

Confidence Intervals for the Difference of Two Means: Independent Samples

Variances assumed to be unequal:

$$(\bar{X}_1 - \bar{X}_2) - t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$\text{d.f.} = \text{smaller value of } n_1 - 1 \text{ or } n_2 - 1$$

EXAMPLE 9–5

Find the 95% confidence interval for the data in Example 9–4.

SOLUTION

Substitute in the formula.

$$(\bar{X}_1 - \bar{X}_2) - t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$(5.5 - 4.3) - 2.447\sqrt{\frac{1.23^2}{9} + \frac{1.19^2}{7}} < \mu_1 - \mu_2 < (5.5 - 4.3) + 2.447\sqrt{\frac{1.23^2}{9} + \frac{1.19^2}{7}}$$

$$1.2 - 1.489 < \mu_1 - \mu_2 < 1.2 + 1.489$$

$$-0.289 < \mu_1 - \mu_2 < 2.689$$

Since 0 is contained in the interval, there is not enough evidence to support the claim that the means are different.

In many statistical software packages, a different method is used to compute the degrees of freedom for this *t* test. They are determined by the formula

$$\text{d.f.} = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1 - 1) + (s_2^2/n_2)^2/(n_2 - 1)}$$

This formula will not be used in this textbook.

There are actually two different options for the use of *t* tests. *One option is used when the variances of the populations are not equal, and the other option is used when the variances are equal.* To determine whether two sample variances are equal, the researcher can use an *F* test, as shown in Section 9–5.

When the variances are assumed to be equal, this formula is used and

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

follows the format of

$$\text{Test value} = \frac{(\text{observed value}) - (\text{expected value})}{\text{standard error}}$$

For the numerator, the terms are the same as in the previously given formula. However, a note of explanation is needed for the denominator of the second test statistic. Since both populations are assumed to have the same variance, the standard error is computed with what is called a pooled estimate of the variance. A **pooled estimate of the variance** is a weighted average of the variance using the two sample variances and the *degrees of freedom* of each variance as the weights. Again, since the algebraic derivation of the standard error is somewhat complicated, it is omitted.

Note, however, that not all statisticians are in agreement about using the *F* test before using the *t* test. Some believe that conducting the *F* and *t* tests at the same level of significance will change the overall level of significance of the *t* test. Their reasons are beyond the scope of this text. Because of this, we will assume that $\sigma_1 \neq \sigma_2$ in this text.

Applying the Concepts 9–2

Too Long on the Telephone

A company collects data on the lengths of telephone calls made by employees in two different divisions. The sample mean and the sample standard deviation for the sales division are 10.26 and 8.56, respectively. The sample mean and sample standard deviation for the shipping and receiving division are 6.93 and 4.93, respectively. A hypothesis test was run, and the computer output follows.

```
Degrees of freedom = 56
Confidence interval limits = −0.18979, 6.84979
Test statistic t = 1.89566
Critical value t = −2.0037, 2.0037
P-value = 0.06317
Significance level = 0.05
```

1. Are the samples independent or dependent?
2. Which number from the output is compared to the significance level to check if the null hypothesis should be rejected?
3. Which number from the output gives the probability of a type I error that is calculated from the sample data?
4. Was a right-, left-, or two-tailed test done? Why?
5. What are your conclusions?
6. What would your conclusions be if the level of significance were initially set at 0.10?

See pages 544–545 for the answers.

Exercises 9–2

For these exercises, perform each of these steps. Assume that all variables are normally or approximately normally distributed.

a. State the hypotheses and identify the claim.
b. Find the critical value(s).
c. Compute the test value.
d. Make the decision.
e. Summarize the results.

Use the traditional method of hypothesis testing unless otherwise specified and assume the variances are unequal.

1. **Waterfall Heights** Is there a significant difference at $\alpha = 0.10$ in the mean heights in feet of waterfalls in Europe and the ones in Asia? The data are shown.

Europe			Asia		
487	1246	1385	614	722	964
470	1312	984	1137	320	830
900	345	820	350	722	1904

Source: World Almanac and Book of Facts.

2. **Tax-Exempt Properties** A tax collector wishes to see if the mean values of the tax-exempt properties are different for two cities. The values of the tax-exempt properties for the two random samples are shown. The data are given in millions of dollars. At $\alpha = 0.05$, is there enough evidence to support the tax collector's claim that the means are different?

City A				City B			
113	22	14	8	82	11	5	15
25	23	23	30	295	50	12	9
44	11	19	7	12	68	81	2
31	19	5	2	20	16	4	5

3. **Noise Levels in Hospitals** The mean noise level of 20 randomly selected areas designated as "casualty doors" was 63.1 dBA, and the sample standard deviation is 4.1 dBA. The mean noise level for 24 randomly selected areas designated as operating theaters was 56.3 dBA, and the sample standard deviation was 7.5 dBA. At $\alpha = 0.05$, can it be concluded that there is a difference in the means?

4. **Ages of Gamblers** The mean age of a random sample of 25 people who were playing the slot machines is 48.7 years, and the standard deviation is 6.8 years. The mean age of a random sample of 35 people who were playing roulette is 55.3 with a standard deviation of 3.2 years. Can it be concluded at $\alpha = 0.05$ that the mean age of those playing the slot machines is less than those playing roulette?

5. **Carbohydrates in Candies** The number of grams of carbohydrates contained in 1-ounce servings of randomly selected chocolate and nonchocolate candy is listed here. Is there sufficient evidence to conclude that the difference in the means is statistically significant? Use $\alpha = 0.10$.

Chocolate:	29	25	17	36	41	25	32	29
	38	34	24	27	29			
Nonchocolate:	41	41	37	29	30	38	39	10
	29	55	29					

Source: The Doctor's Pocket Calorie, Fat, and Carbohydrate Counter.

6. **Weights of Vacuum Cleaners** Upright vacuum cleaners have either a hard body type or a soft body type. Shown are the weights in pounds of a random sample of each type. At $\alpha = 0.05$, can it be concluded that the means of the weights are different?

Hard body types				Soft body types			
21	17	17	20	24	13	11	13
16	17	15	20	12	15		
23	16	17	17				
13	15	16	18				
18							

7. **Weights of Running Shoes** The weights in ounces of a sample of running shoes for men and women are shown. Test the claim that the means are different. Use the *P*-value method with $\alpha = 0.05$.

Men		Women		
10.4	12.6	10.6	10.2	8.8
11.1	14.7	9.6	9.5	9.5
10.8	12.9	10.1	11.2	9.3
11.7	13.3	9.4	10.3	9.5
12.8	14.5	9.8	10.3	11.0

8. **Teacher Salaries** A researcher claims that the mean of the salaries of elementary school teachers is greater than the mean of the salaries of secondary school teachers in a large school district. The mean of the salaries of a random sample of 26 elementary school teachers is \$48,256, and the sample standard deviation is \$3,912.40. The mean of the salaries of a random sample of 24 secondary school teachers is \$45,633. The sample standard deviation is \$5533. At $\alpha = 0.05$, can it be concluded that the mean of the salaries of the elementary school teachers is greater than the mean of the salaries of the secondary school teachers? Use the *P*-value method.

9. Find the 90% confidence for the difference of the means in Exercise 1 of this section.

10. Find the 95% confidence interval for the difference of the means in Exercise 6 of this section.

11. **Hours Spent Watching Television** According to Nielsen Media Research, children (ages 2–11) spend an average of 21 hours 30 minutes watching television per week while teens (ages 12–17) spend an average of 20 hours 40 minutes. Based on the sample statistics shown, is there sufficient evidence to conclude a

difference in average television watching times between the two groups? Use $\alpha = 0.01$.

	Children	**Teens**
Sample mean	22.45	18.50
Sample variance	16.4	18.2
Sample size	15	15

Source: Time Almanac.

12. **Professional Golfers' Earnings** Two random samples of earnings of professional golfers were selected. One sample was taken from the Professional Golfers Association, and the other was taken from the Ladies Professional Golfers Association. At $\alpha = 0.05$, is there a difference in the means? The data are in thousands of dollars.

PGA

446	1147	1344	9188	5687
10,508	4910	8553	7573	375

LPGA

48	76	122	466	863
100	1876	2029	4364	2921

13. **Cyber School Enrollment** The data show the number of students attending cyber charter schools in Allegheny County and the number of students attending cyber schools in counties surrounding Allegheny County. At $\alpha = 0.01$, is there enough evidence to support the claim that the average number of students in school districts in Allegheny County who attend cyber schools is greater than those who attend cyber schools in school districts outside Allegheny County? Give a factor that should be considered in interpreting this answer.

Allegheny County						**Outside Allegheny County**					
25	75	38	41	27	32	57	25	38	14	10	29

Source: Pittsburgh Tribune-Review.

14. **Hockey's Highest Scorers** The number of points held by random samples of the NHL's highest scorers for both the Eastern Conference and the Western Conference is shown. At $\alpha = 0.05$, can it be concluded that there is a difference in means based on these data?

Eastern Conference				**Western Conference**			
83	60	75	58	77	59	72	58
78	59	70	58	37	57	66	55
62	61	59		61			

Source: www.foxsports.com

15. **Hospital Stays for Maternity Patients** Health Care Knowledge Systems reported that an insured woman spends on average 2.3 days in the hospital for a routine childbirth, while an uninsured woman spends on average 1.9 days. Assume two random samples of 16 women each were used in both samples. The standard deviation of the first sample is equal to 0.6 day, and the standard deviation of the second sample is 0.3 day. At $\alpha = 0.01$, test the claim that the means are equal. Find the 99% confidence interval for the differences of the means. Use the P-value method.

Source: Michael D. Shook and Robert L. Shook, *The Book of Odds.*

16. **Ages of Homes** Whiting, Indiana, leads the "Top 100 Cities with the Oldest Houses" list with the average age of houses being 66.4 years. Farther down the list resides Franklin, Pennsylvania, with an average house age of 59.4 years. Researchers selected a random sample of 20 houses in each city and obtained the following statistics. At $\alpha = 0.05$, can it be concluded that the houses in Whiting are older? Use the P-value method.

	Whiting	**Franklin**
Mean age	62.1 years	55.6 years
Standard deviation	5.4 years	3.9 years

Source: www.city-data.com

17. **Medical School Enrollments** A random sample of enrollments from medical schools that specialize in research and from those that are noted for primary care is listed. Find the 90% confidence interval for the difference in the means.

Research				**Primary care**			
474	577	605	663	783	605	427	728
783	467	670	414	546	474	371	107
813	443	565	696	442	587	293	277
692	694	277	419	662	555	527	320
884							

Source: U.S. News & World Report Best Graduate Schools.

18. **Out-of-State Tuitions** The out-of-state tuitions (in dollars) for random samples of both public and private four-year colleges in a New England state are listed. Find the 95% confidence interval for the difference in the means.

Private		**Public**	
13,600	13,495	7,050	9,000
16,590	17,300	6,450	9,758
23,400	12,500	7,050	7,871
		16,100	

Source: New York Times Almanac.

19. **Gasoline Prices** A random sample of monthly gasoline prices was taken from 2011 and from 2015. The samples are shown. Using $\alpha = 0.01$, can it be concluded that gasoline cost more in 2015? Use the P-value method.

2011	2.02	2.47	2.50	2.70	3.13	2.56	
2015	2.36	2.46	2.63	2.76	3.00	2.85	2.77

20. **Miniature Golf Scores** A large group of friends went miniature golfing together at a par 54 course and decided to play on two teams. A random sample of scores from each of the two teams is shown. At $\alpha = 0.05$, is there a difference in mean scores between the two teams? Use the P-value method.

Team 1	61	44	52	47	56	63	62	55
Team 2	56	40	42	58	48	52	51	

21. Home Runs Two random samples of professional baseball players were selected and the number of home runs hit were recorded. One sample was obtained from the National League, and the other sample was obtained from the American League. At $\alpha = 0.10$, is there a difference in the means?

National League				American League			
18	4	8	2	6	11	18	11
9	2	6	5	3	12	25	4
6	8	29	25	24	9	12	5

22. Batting Averages Random samples of batting averages from the leaders in both leagues prior to the All-Star break are shown. At the 0.05 level of significance, can a difference be concluded?

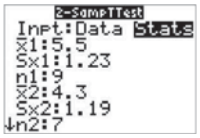
National	.360	.654	.652	.338	.313	.309
American	.340	.332	.317	.316	.314	.306

Technology Step by Step

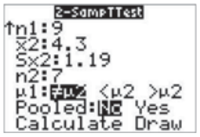
TI-84 Plus Step by Step

Hypothesis Test for the Difference Between Two Means and *t* Distribution (Statistics)

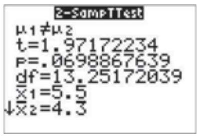
This refers to Example 9–4 in the text.



```
2-SampTTest
Inpt:Data Stats
x̄1:5.5
Sx1:1.23
n1:9
x̄2:4.3
Sx2:1.19
↓n2:7
```



```
2-SampTTest
↑n1:9
 x̄2:4.3
 Sx2:1.19
 n2:7
 μ1:≠μ2 <μ2 >μ2
 Pooled:No Yes
 Calculate Draw
```



```
2-SampTTest
 μ1≠μ2
 t=1.97172234
 p=.0698867639
 df=13.25172039
 x̄1=5.5
↓x̄2=4.3
```

Example TI9–4

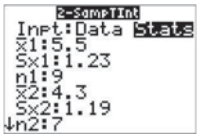
1. Press **STAT** and move the cursor to TESTS.
2. Press **4** for 2-SampTTest.
3. Move the cursor to Stats and press **ENTER.**
4. Type in the appropriate values.
5. Move the cursor to the appropriate alternative hypothesis and press **ENTER.**
6. On the line for Pooled, move the cursor to No (standard deviations are assumed not equal) and press **ENTER.**
7. Move the cursor to Calculate and press **ENTER.**

Confidence Interval for the Difference Between Two Means and *t* Distribution (Data)

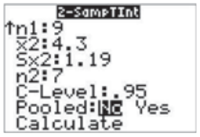
1. Enter the data values into L_1 and L_2.
2. Press **STAT** and move the cursor to TESTS.
3. Press **0** for 2-SampTInt.
4. Move the cursor to Data and press **ENTER.**
5. Type in the appropriate values.
6. On the line for Pooled, move the cursor to No (standard deviations are assumed not equal) and press **ENTER.**
7. Move the cursor to Calculate and press **ENTER.**

Confidence Interval for the Difference Between Two Means and *t* Distribution (Statistics)

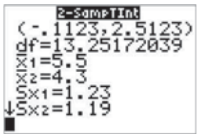
This refers to Example 9–5 in the text.



```
2-SampTInt
Inpt:Data Stats
x̄1:5.5
Sx1:1.23
n1:9
x̄2:4.3
Sx2:1.19
↓n2:7
```



```
2-SampTInt
↑n1:9
 x̄2:4.3
 Sx2:1.19
 n2:7
 C-Level:.95
 Pooled:No Yes
 Calculate
```



```
2-SampTInt
 (-.1123,2.5123)
 df=13.25172039
 x̄1=5.5
 x̄2=4.3
 Sx1=1.23
↓Sx2=1.19
```

Example TI9–5

1. Press **STAT** and move the cursor to TESTS.
2. Press **0** for 2-SampTInt.
3. Move the cursor to Stats and press **ENTER.**
4. Type in the appropriate values.
5. On the line for Pooled, move the cursor to No (standard deviations are assumed not equal) and press **ENTER.**
6. Move the cursor to Calculate and press **ENTER.**

EXCEL Step by Step

Testing the Difference Between Two Means: Independent Samples

Excel has a two-sample *t* test included in the Data Analysis Add-in. The following example shows how to perform a *t* test for the difference between two means.

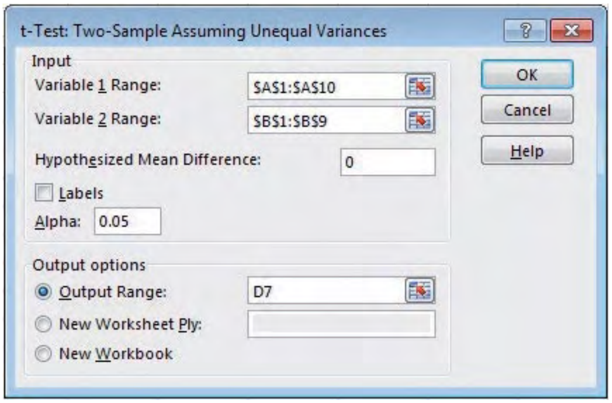
Example XL9–2

Test the claim that there is no difference between population means based on these sample data. Assume the population variances are not equal. Use $\alpha = 0.05$.

Set A	32	38	37	36	36	34	39	36	37	42
Set B	30	36	35	36	31	34	37	33	32	

1. Enter the 10-number data set A into column A.
2. Under the Home tab, select Format > enter the 9-number data set B into column B.
3. Select the Data tab from the toolbar. Then select Data Analysis.
4. In the Data Analysis box, under Analysis Tools select *t*-test: Two-Sample Assuming Unequal Variances, and click [OK].
5. In Input, type in the Variable 1 Range: **A1:A10** and the Variable 2 Range: **B1:B9.**
6. Type **0** for the Hypothesized Mean Difference.
7. Type **0.05** for Alpha.
8. In Output options, type D7 for the Output Range, then click [OK].

Two-Sample *t* Test in Excel



t-Test: Two-Sample Assuming Unequal Variances

Input

Variable 1 Range: $A$1:$A$10

Variable 2 Range: $B$1:$B$9

Hypothesized Mean Difference: 0

☐ Labels

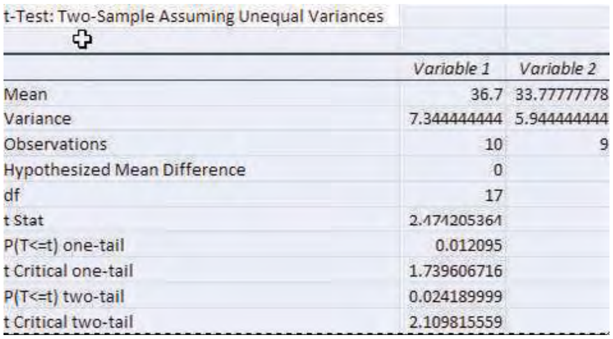
Alpha: 0.05

Output options

● Output Range: D7

○ New Worksheet Ply:

○ New Workbook



t-Test: Two-Sample Assuming Unequal Variances

	Variable 1	*Variable 2*
Mean	36.7	33.77777778
Variance	7.344444444	5.944444444
Observations	10	9
Hypothesized Mean Difference	0	
df	17	
t Stat	2.474205364	
P(T<=t) one-tail	0.012095	
t Critical one-tail	1.739606716	
P(T<=t) two-tail	0.024189999	
t Critical two-tail	2.109815559	

Note: You may need to increase the column width to see all the results. To do this:

1. Highlight the columns D, E, and F.
2. Select **Format>AutoFit** Column Width.

The output reports both one- and two-tailed *P*-values.

MINITAB
Step by Step

Test the Difference Between Two Means: Independent Samples*

MINITAB will calculate the test statistic and *P*-value for differences between the means for two populations when the population standard deviations are unknown.

For Example 9–2, is the average number of sports for men higher than the average number for women?

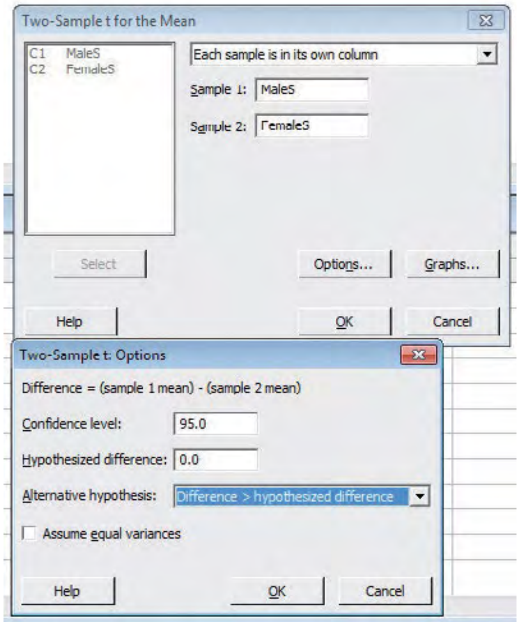
1. Enter the data for Example 9–2 into C1 and C2. Name the columns **MaleS** and **FemaleS.**
2. Select **Stat>Basic Statistics>2-Sample t.**
3. Select Each sample is in its own column from the drop down menu.

---

*MINITAB does not calculate a z test statistic. This statistic can be used instead.

There is one sample in each column.

4. Click in the box for Sample 1. Double-click C1 MaleS in the list.
5. Click in the box for Sample 2, then double-click C2 FemaleS in the list. Do not check the box for Assume equal variances. MINITAB will use the large sample formula. The completed dialog box is shown.
6. Click [Options].
   a) Type in **90** for the Confidence level and **0** for the Hypothesized difference.
   b) Select Difference > hypothesized difference for the Alternative hypothesis. Make sure that Assume equal variances is not checked.
7. Click [OK] twice. Since the *P*-value is greater than the significance level, 0.172 > 0.1, do not reject the null hypothesis.



Two-Sample t for the Mean

C1 MaleS
C2 FemaleS

Each sample is in its own column

Sample 1: MaleS

Sample 2: FemaleS

Two-Sample t: Options

Difference = (sample 1 mean) - (sample 2 mean)

Confidence level: 95.0

Hypothesized difference: 0.0

Alternative hypothesis: Difference > hypothesized difference

☐ Assume equal variances

Two-Sample T-Test and CI: MaleS, FemaleS

```
Two–sample  T  for  MaleS  vs FemaleS

              N      Mean      StDev      SE Mean
MaleS        50      8.56       3.26         0.46
FemaleS      50      7.94       3.27         0.46

Difference = μ  (MaleS) – μ (FemaleS)
Estimate for difference: 0.620
95% lower bound for difference: –0.464
T-Test of difference = 0 (vs >): T–Value = 0.95  P-Value = 0.172  DF = 97
```

9–3 Testing the Difference Between Two Means: Dependent Samples

OBJECTIVE 3

Test the difference between two means for dependent samples.

In Section 9–1, the z test was used to compare two sample means when the samples were independent and σ_1 and σ_2 were known. In Section 9–2, the t test was used to compare two sample means when the samples were independent. In this section, a different version of the t test is explained. This version is used when the samples are dependent. Samples are considered to be **dependent samples** when the subjects are paired or matched in some way. Dependent samples are sometimes called matched-pair samples.

For example, suppose a medical researcher wants to see whether a drug will affect the reaction time of its users. To test this hypothesis, the researcher must pretest the subjects in the sample. That is, they are given a test to ascertain their normal reaction times. Then after taking the drug, the subjects are tested again, using a posttest. Finally, the means of the two tests are compared to see whether there is a difference. Since the same subjects are used in both cases, the samples are *related;* subjects scoring high on the pretest will generally score high on the posttest, even after consuming the drug. Likewise, those scoring lower on the pretest will tend to score lower on the posttest. To take this effect into account, the researcher employs a t test, using the differences between the pretest values and the posttest values. Thus, only the gain or loss in values is compared.

Here are some other examples of dependent samples. A researcher may want to design an SAT preparation course to help students raise their test scores the second time they take the SAT. Hence, the differences between the two exams are compared. A medical specialist may want to see whether a new counseling program will help subjects lose weight. Therefore, the preweights of the subjects will be compared with the postweights.

Besides samples in which the same subjects are used in a pre-post situation, there are other cases where the samples are considered dependent. For example, students might be matched or paired according to some variable that is pertinent to the study; then one student is assigned to one group, and the other student is assigned to a second group. For instance, in a study involving learning, students can be selected and paired according to their IQs. That is, two students with the same IQ will be paired. Then one will be assigned to one sample group (which might receive instruction by computers), and the other student will be assigned to another sample group (which might receive instruction by the lecture discussion method). These assignments will be done randomly. Since a student's IQ is important to learning, it is a variable that should be controlled. By matching subjects on IQ, the researcher can eliminate the variable's influence, for the most part. Matching, then, helps to reduce type II error by eliminating extraneous variables.

Two notes of caution should be mentioned. First, when subjects are matched according to one variable, the matching process does not eliminate the influence of other variables. Matching students according to IQ does not account for their mathematical ability or their familiarity with computers. Since not all variables influencing a study can be controlled, it is up to the researcher to determine which variables should be used in matching. Second, when the same subjects are used for a pre-post study, sometimes the knowledge that they are participating in a study can influence the results. For example, if people are placed in a special program, they may be more highly motivated to succeed simply because they have been selected to participate; the program itself may have little effect on their success.

When the samples are dependent, a special *t* test for dependent means is used. This test employs the difference in values of the matched pairs. The hypotheses are as follows:

Two-tailed	Left-tailed	Right-tailed
H_0: $\mu_D = 0$ H_1: $\mu_D \neq 0$	H_0: $\mu_D = 0$ H_1: $\mu_D < 0$	H_0: $\mu_D = 0$ H_1: $\mu_D > 0$

Here, μ_D is the symbol for the expected mean of the difference of the matched pairs. The general procedure for finding the test value involves several steps.

First, find the differences of the values of the pairs of data.

$$D = X_1 - X_2$$

Second, find the mean $\overline{D}$ of the differences, using the formula

$$\overline{D} = \frac{\Sigma D}{n}$$

where n is the number of data pairs. Third, find the standard deviation s_D of the differences, using the formula

$$s_D = \sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{n(n-1)}}$$

Fourth, find the estimated standard error $s_{\overline{D}}$ of the differences, which is

$$s_{\overline{D}} = \frac{s_D}{\sqrt{n}}$$

Finally, find the test value, using the formula

$$t = \frac{\overline{D} - \mu_D}{s_D/\sqrt{n}} \qquad \text{with d.f.} = n - 1$$

The formula in the final step follows the basic format of

$$\text{Test value} = \frac{\text{(observed value)} - \text{(expected value)}}{\text{standard error}}$$

where the observed value is the mean of the differences. The expected value μ_D is zero if the hypothesis is $\mu_D = 0$. The standard error of the difference is the standard deviation of the difference, divided by the square root of the sample size. Both populations must be normally or approximately normally distributed.

Before you can use the testing method presented in this section, the following assumptions must be met.

Assumptions for the *t* Test for Two Means When the Samples Are Dependent

1. The sample or samples are random.
2. The sample data are dependent.
3. When the sample size or sample sizes are less than 30, the population or populations must be normally or approximately normally distributed.

In this book, the assumptions will be stated in the exercises; however, when encountering statistics in other situations, you must check to see that these assumptions have been met before proceeding.

The formulas for this *t* test are given next.

Formulas for the *t* Test for Dependent Samples

$$t = \frac{\overline{D} - \mu_D}{s_D/\sqrt{n}}$$

with d.f. $= n - 1$ and where

$$\overline{D} = \frac{\Sigma D}{n} \qquad \text{and} \qquad s_D = \sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{n(n-1)}}$$

The steps for this *t* test are summarized in the Procedure Table.

Procedure Table

Testing the Difference Between Means for Dependent Samples

Step 1 State the hypotheses and identify the claim.

Step 2 Find the critical value(s).

Step 3 Compute the test value.

a. Make a table, as shown.

X_1	X_2	A $D = X_1 - X_2$	B $D^2 = (X_1 - X_2)^2$
⋮	⋮	$\Sigma D =$ ______	$\Sigma D^2 =$ ______

b. Find the differences and place the results in column A.

$$D = X_1 - X_2$$

c. Find the mean of the differences.

$$\overline{D} = \frac{\Sigma D}{n}$$

d. Square the differences and place the results in column B. Complete the table.

$$D^2 = (X_1 - X_2)^2$$

e. Find the standard deviation of the differences.

$$s_D = \sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{n(n-1)}}$$

f. Find the test value.

$$t = \frac{\overline{D} - \mu_D}{s_D/\sqrt{n}} \qquad \text{with d.f.} = n - 1$$

Step 4 Make the decision.

Step 5 Summarize the results.

Unusual Stat

About 4% of Americans spend at least one night in jail each year.

EXAMPLE 9–6 Bank Deposits

A random sample of nine local banks shows their deposits (in billions of dollars) 3 years ago and their deposits (in billions of dollars) today. At $\alpha = 0.05$, can it be concluded that the average in deposits for the banks is greater today than it was 3 years ago? Use $\alpha = 0.05$. Assume the variable is normally distributed.

Source: SNL Financial.

Bank	1	2	3	4	5	6	7	8	9
3 years ago	11.42	8.41	3.98	7.37	2.28	1.10	1.00	0.9	1.35
Today	16.69	9.44	6.53	5.58	2.92	1.88	1.78	1.5	1.22

SOLUTION

Step 1 State the hypothesis and identify the claim. Since we are interested to see if there has been an increase in deposits, the deposits 3 years ago must be less than the deposits today; hence, the deposits must be significantly less 3 years ago than they are today. Hence, the mean of the differences must be less than zero.

$$H_0\colon \mu_D = 0 \qquad \text{and} \qquad H_1\colon \mu_D < 0 \text{ (claim)}$$

Step 2 Find the critical value. The degrees of freedom are $n - 1$, or $9 - 1 = 8$. Using Table F, the critical value for a left-tailed test with $\alpha = 0.05$ is -1.860.

Step 3 Compute the test value.

a. Make a table.

3 years ago (X_1)	Today (X_2)	A $D = X_1 - X_2$	B $D^2 = (X_1 - X_2)^2$
11.42	16.69		
8.41	9.44		
3.98	6.53		
7.37	5.58		
2.28	2.92		
1.10	1.88		
1.00	1.78		
0.90	1.50		
1.35	1.22		

b. Find the differences and place the results in column A.

$$\begin{aligned}
11.42 - 16.69 &= -5.27\\
8.41 - 9.44 &= -1.03\\
3.98 - 6.53 &= -2.55\\
7.37 - 5.58 &= +1.79\\
2.28 - 2.92 &= -0.64\\
1.10 - 1.88 &= -0.78\\
1.00 - 1.78 &= -0.78\\
0.9 - 1.50 &= -0.60\\
1.35 - 1.22 &= \underline{+0.13}\\
\Sigma D &= -9.73
\end{aligned}$$

c. Find the means of the differences.

$$\overline{D} = \frac{\Sigma D}{n} = \frac{-9.73}{9} = -1.081$$

d. Square the differences and place the results in column B.

$$\begin{aligned}(-5.27)^2 &= 27.7729\\(-1.03)^2 &= 1.0609\\(-2.55)^2 &= 6.5025\\(+1.79)^2 &= 3.2041\\(-0.64)^2 &= 0.4096\\(-0.78)^2 &= 0.6084\\(-0.78)^2 &= 0.6084\\(-0.60)^2 &= 0.3600\\(+0.13)^2 &= \underline{0.0169}\\\Sigma D^2 &= 40.5437\end{aligned}$$

The completed table is shown next.

3 years ago (X_1)	Today (X_2)	A $D = X_1 - X_2$	B $D^2 = (X_1 - X_2)^2$
11.42	16.69	−5.27	27.7729
8.41	9.44	−1.03	1.0609
3.98	6.53	−2.55	6.5025
7.37	5.58	+1.79	3.2041
2.28	2.92	−0.64	0.4096
1.10	1.88	−0.78	0.6084
1.00	1.78	−0.78	0.6084
0.90	1.50	−0.60	0.3600
1.35	1.22	+0.13	0.0169
		$\Sigma D = -9.73$	$\Sigma D^2 = 40.5437$

e. Find the standard deviation of the differences.

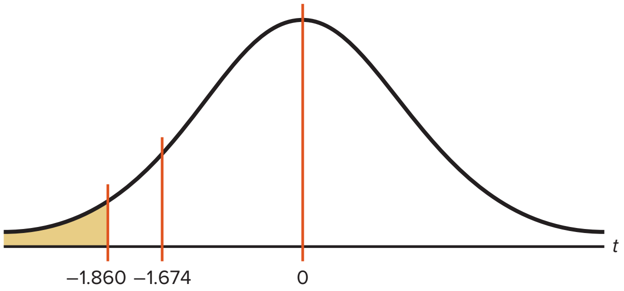
$$\begin{aligned}s_D &= \sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{n(n-1)}}\\&= \sqrt{\frac{9(40.5437) - (-9.73)^2}{9(9-1)}}\\&= \sqrt{\frac{270.2204}{72}}\\&= 1.937\end{aligned}$$

f. Find the test value.

$$t = \frac{\overline{D} - \mu_D}{s_D/\sqrt{n}} = \frac{-1.081 - 0}{1.937/\sqrt{9}} = -1.674$$

Step 4 Make the decision. Do not reject the null hypothesis since the test value, −1.674, is greater than the critical value, −1.860. See Figure 9–6.

FIGURE 9–6 Critical and Test Values for Example 9–6



Step 5 Summarize the results. There is not enough evidence to show that the deposits have increased over the last 3 years.

EXAMPLE 9–7 Cholesterol Levels

A dietitian wishes to see if a person's cholesterol level will change if the diet is supplemented by a certain mineral. Six randomly selected subjects were pretested, and then they took the mineral supplement for a 6-week period. The results are shown in the table. (Cholesterol level is measured in milligrams per deciliter.) Can it be concluded that the cholesterol level has been changed at $\alpha = 0.10$? Assume the variable is approximately normally distributed.

Subject	1	2	3	4	5	6
Before (X_1)	210	235	208	190	172	244
After (X_2)	190	170	210	188	173	228

SOLUTION

Step 1 State the hypotheses and identify the claim. If the diet is effective, the before cholesterol levels should be different from the after levels.

$$H_0\colon \mu_D = 0 \qquad \text{and} \qquad H_1\colon \mu_D \neq 0 \text{ (claim)}$$

Step 2 Find the critical value. The degrees of freedom are $6 - 1 = 5$. At $\alpha = 0.10$, the critical values are ± 2.015.

Step 3 Compute the test value.

a. Make a table.

Before (X_1)	**After (X_2)**	**A** $D = X_1 - X_2$	**B** $D^2 = (X_1 - X_2)^2$
210	190		
235	170		
208	210		
190	188		
172	173		
244	228		

b. Find the differences and place the results in column A.

$$\begin{aligned}
210 - 190 &= 20 \\
235 - 170 &= 65 \\
208 - 210 &= -2 \\
190 - 188 &= 2 \\
172 - 173 &= -1 \\
244 - 228 &= \underline{16} \\
\Sigma D &= 100
\end{aligned}$$

c. Find the mean of the differences.

$$\overline{D} = \frac{\Sigma D}{n} = \frac{100}{6} = 16.7$$

d. Square the differences and place the results in column B.

$$\begin{aligned}
(20)^2 &= 400 \\
(65)^2 &= 4225 \\
(-2)^2 &= 4 \\
(2)^2 &= 4 \\
(-1)^2 &= 1 \\
(16)^2 &= \underline{256} \\
\Sigma D^2 &= 4890
\end{aligned}$$

Then complete the table as shown.

Before (X_1)	After (X_2)	A $D = X_1 - X_2$	B $D^2 = (X_1 - X_2)^2$
210	190	20	400
235	170	65	4225
208	210	–2	4
190	188	2	4
172	173	–1	1
244	228	16	256
		$\Sigma D = 100$	$\Sigma D^2 = 4890$

e. Find the standard deviation of the differences.

$$s_D = \sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{n(n-1)}}$$

$$= \sqrt{\frac{6 \cdot 4890 - 100^2}{6(6-1)}}$$

$$= \sqrt{\frac{29{,}340 - 10{,}000}{30}}$$

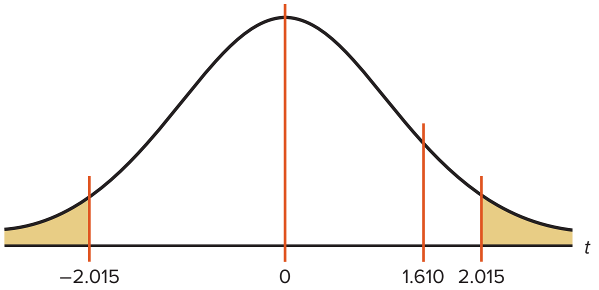
$$= 25.4$$

f. Find the test value.

$$t = \frac{\overline{D} - \mu_D}{s_D/\sqrt{n}} = \frac{16.7 - 0}{25.4/\sqrt{6}} = 1.610$$

Step 4 Make the decision. The decision is to not reject the null hypothesis, since the test value 1.610 is in the noncritical region, as shown in Figure 9–7.

FIGURE 9–7 Critical and Test Values for Example 9–7



Step 5 Summarize the results. There is not enough evidence to support the claim that the mineral changes a person's cholesterol level.

The *P*-values for the *t* test are found in Table F. For a two-tailed test with d.f. = 5 and $t = 1.610$, the *P*-value is found between 1.476 and 2.015; hence, $0.10 < P\text{-value} < 0.20$. Thus, the null hypothesis cannot be rejected at $\alpha = 0.10$.

If a specific difference is hypothesized, this formula should be used

$$t = \frac{\overline{D} - \mu_D}{s_D/\sqrt{n}}$$

where μ_D is the hypothesized difference.

SPEAKING OF STATISTICS Can Video Games Save Lives?

Can playing video games help doctors perform surgery? The answer is yes. A study showed that surgeons who played video games for at least 3 hours each week made about 37% fewer mistakes and finished operations 27% faster than those who did not play video games.

The type of surgery that they performed is called *laparoscopic* surgery, where the surgeon inserts a tiny video camera into the body and uses a joystick to maneuver the surgical instruments while watching the results on a television monitor. This study compares two groups and uses proportions. What statistical test do you think was used to compare the percentages? (See Section 9–4.)

© Antonia Reeve/Science Source

For example, if a dietitian claims that people on a specific diet will lose an average of 3 pounds in a week, the hypotheses are

$$H_0\colon \mu_D = 3 \qquad \text{and} \qquad H_1\colon \mu_D \neq 3$$

The value 3 will be substituted in the test statistic formula for μ_D.

Confidence intervals can be found for the mean differences with this formula.

Confidence Interval for the Mean Difference

$$\overline{D} - t_{\alpha/2}\frac{s_D}{\sqrt{n}} < \mu_D < \overline{D} + t_{\alpha/2}\frac{s_D}{\sqrt{n}}$$

$$\text{d.f.} = n - 1$$

EXAMPLE 9–8

Find the 90% confidence interval for the data in Example 9–7.

SOLUTION

Substitute in the formula.

$$\overline{D} - t_{\alpha/2}\frac{s_D}{\sqrt{n}} < \mu_D < \overline{D} + t_{\alpha/2}\frac{s_D}{\sqrt{n}}$$

$$16.7 - 2.015 \cdot \frac{25.4}{\sqrt{6}} < \mu_D < 16.7 + 2.015 \cdot \frac{25.4}{\sqrt{6}}$$

$$16.7 - 20.89 < \mu_D < 16.7 + 20.89$$

$$-4.19 < \mu_D < 37.59$$

$$-4.2 < \mu_D < 37.6$$

Since 0 is contained in the interval, the decision is to not reject the null hypothesis $H_0\colon \mu_D = 0$. Hence, there is not enough evidence to support the claim that the mineral changes a person's cholesterol, as previously shown.

Applying the Concepts 9–3

Air Quality

As a researcher for the EPA, you have been asked to determine if the air quality in the United States has changed over the past 2 years. You select a random sample of 10 metropolitan areas and find the number of days each year that the areas failed to meet acceptable air quality standards. The data are shown.

Year 1	18	125	9	22	138	29	1	19	17	31
Year 2	24	152	13	21	152	23	6	31	34	20

Source: The World Almanac and Book of Facts.

Based on the data, answer the following questions.

1. What is the purpose of the study?
2. Are the samples independent or dependent?
3. What hypotheses would you use?
4. What is (are) the critical value(s) that you would use?
5. What statistical test would you use?
6. How many degrees of freedom are there?
7. What is your conclusion?
8. Could an independent means test have been used?
9. Do you think this was a good way to answer the original question?

See page 545 for the answers.

Exercises 9–3

1. Classify each as independent or dependent samples.
   - *a.* Heights of identical twins
   - *b.* Test scores of the same students in English and psychology
   - *c.* The effectiveness of two different brands of aspirin on two different groups of people
   - *d.* Effects of a drug on reaction time of two different groups of people, measured by a before-and-after test
   - *e.* The effectiveness of two different diets on two different groups of individuals

For Exercises 2 through 12, perform each of these steps. Assume that all variables are normally or approximately normally distributed.

- *a.* State the hypotheses and identify the claim.
- *b.* Find the critical value(s).
- *c.* Compute the test value.
- *d.* Make the decision.
- *e.* Summarize the results.

Use the traditional method of hypothesis testing unless otherwise specified.

2. **Retention Test Scores** A random sample of non-English majors at a selected college was used in a study to see if the student retained more from reading a 19th-century novel or by watching it in DVD form. Each student was assigned one novel to read and a different one to watch, and then they were given a 100-point written quiz on each novel. The test results are shown. At $\alpha = 0.05$, can it be concluded that the book scores are higher than the DVD scores?

Book	90	80	90	75	80	90	84
DVD	85	72	80	80	70	75	80

3. **Improving Study Habits** As an aid for improving students' study habits, nine students were randomly selected to attend a seminar on the importance of education in life. The table shows the number of hours each student studied per week before and after

the seminar. At $\alpha = 0.10$, did attending the seminar increase the number of hours the students studied per week?

Before	9	12	6	15	3	18	10	13	7
After	9	17	9	20	2	21	15	22	6

4. **Obstacle Course Times** An obstacle course was set up on a campus, and 8 randomly selected volunteers were given a chance to complete it while they were being timed. They then sampled a new energy drink and were given the opportunity to run the course again. The "before" and "after" times in seconds are shown. Is there sufficient evidence at $\alpha = 0.05$ to conclude that the students did better the second time? Discuss possible reasons for your results.

Student	1	2	3	4	5	6	7	8
Before	67	72	80	70	78	82	69	75
After	68	70	76	65	75	78	65	68

5. **Cholesterol Levels** A medical researcher wishes to see if he can lower the cholesterol levels through diet in 6 people by showing a film about the effects of high cholesterol levels. The data are shown. At $\alpha = 0.05$, did the cholesterol level decrease on average?

Patient	1	2	3	4	5	6
Before	243	216	214	222	206	219
After	215	202	198	195	204	213

6. **PGA Golf Scores** At a recent PGA tournament (the Honda Classic at Palm Beach Gardens, Florida) the following scores were posted for eight randomly selected golfers for two consecutive days. At $\alpha = 0.05$, is there evidence of a difference in mean scores for the two days?

Golfer	1	2	3	4	5	6	7	8
Thursday	67	65	68	68	68	70	69	70
Friday	68	70	69	71	72	69	70	70

Source: Washington Observer-Reporter.

7. **Reducing Errors in Grammar** A composition teacher wishes to see whether a new smartphone app will reduce the number of grammatical errors her students make when writing a two-page essay. She randomly selects six students, and the data are shown. At $\alpha = 0.025$, can it be concluded that the number of errors has been reduced?

Student	1	2	3	4	5	6
Errors before	12	9	0	5	4	3
Errors after	9	6	1	3	2	3

8. **Overweight Dogs** A veterinary nutritionist developed a diet for overweight dogs. The total volume of food consumed remains the same, but one-half of the dog food is replaced with a low-calorie "filler" such as canned green beans. Six overweight dogs were randomly selected from her practice and were put on this program. Their initial weights were recorded, and they were weighed again after 4 weeks. At the 0.05 level of significance, can it be concluded that the dogs lost weight?

Before	42	53	48	65	40	52
After	39	45	40	58	42	47

9. **Pulse Rates of Identical Twins** A researcher wanted to compare the pulse rates of identical twins to see whether there was any difference. Eight sets of twins were randomly selected. The rates are given in the table as number of beats per minute. At $\alpha = 0.01$, is there a significant difference in the average pulse rates of twins? Use the *P*-value method. Find the 99% confidence interval for the difference of the two.

Twin A	87	92	78	83	88	90	84	93
Twin B	83	95	79	83	86	93	80	86

10. **Toy Assembly Test** An educational researcher devised a wooden toy assembly project to test learning in 6-year-olds. The time in seconds to assemble the project was noted, and the toy was disassembled out of the child's sight. Then the child was given the task to repeat. The researcher would conclude that learning occurred if the mean of the second assembly times was less than the mean of the first assembly times. At $\alpha = 0.01$, can it be concluded that learning took place? Use the *P*-value method, and find the 99% confidence interval of the difference in means.

Child	1	2	3	4	5	6	7
Trial 1	100	150	150	110	130	120	118
Trial 2	90	130	150	90	105	110	120

11. **Reducing Errors in Spelling** A ninth-grade teacher wishes to see if a new spelling program will reduce the spelling errors in his students' writing. The number of spelling errors made by the students in a five-page report before the program is shown. Then the number of spelling errors made by students in a five-page report after the program is shown. At $\alpha = 0.05$, did the program work?

Before	8	3	10	5	9	11	12
After	6	4	8	1	4	7	11

12. **Mistakes in a Song** A random sample of six music students played a short song, and the number of mistakes in music each student made was recorded. After they practiced the song 5 times, the number of mistakes each student made was recorded. The data are shown. At $\alpha = 0.05$, can it be concluded that there was a decrease in the mean number of mistakes?

Student	A	B	C	D	E	F
Before	10	6	8	8	13	8
After	4	2	2	7	8	9

Extending the Concepts

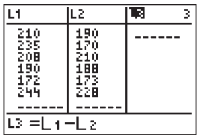
13. Instead of finding the mean of the differences between X_1 and X_2 by subtracting $X_1 - X_2$, you can find it by finding the means of X_1 and X_2 and then subtracting the means. Show that these two procedures will yield the same results.

Technology Step by Step

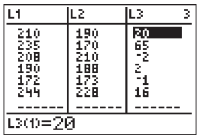
TI-84 Plus Step by Step

Hypothesis Test for the Difference Between Two Means: Dependent Samples

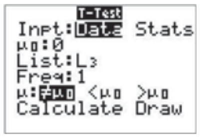
This refers to Example 9–7 in the text.



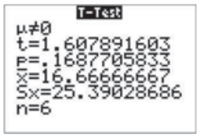
```
L1      L2      L3     3
 210     190    ------
 235     170
 208     210
 190     188
 172     173
 244     228
 ------  ------
L3 =L1-L2
```



```
L1      L2      L3     3
 210     190    20
 235     170    65
 208     210    -2
 190     188    2
 172     173    -1
 244     228    16
 ------  ------ ------
L3(1)=20
```



```
T-Test
Inpt:Data Stats
μ0:0
List:L3
Freq:1
μ:≠μ0 <μ0 >μ0
Calculate Draw
```



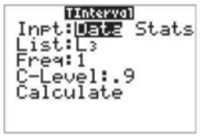
```
T-Test
μ≠0
t=1.607891603
p=.1687705833
x̄=16.66666667
Sx=25.39028686
n=6
```

Example TI9–6

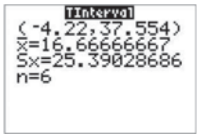
1. Enter the data values into L_1 and L_2.
2. Move the cursor to the top of the L_3 column so that L_3 is highlighted.
3. Type **$L_1 - L_2$,** then press **ENTER.**
4. Press **STAT** and move the cursor to TESTS.
5. Press **2** for TTest.
6. Move the cursor to Data and press **ENTER.**
7. Type in the appropriate values, using **0** for μ_0 and **L_3** for the list.
8. Move the cursor to the appropriate alternative hypothesis and press **ENTER.**
9. Move the cursor to Calculate and press **ENTER.**

Confidence Interval for the Difference Between Two Means: Dependent Samples

This refers to Example 9–8 in the text.



```
TInterval
Inpt:Data Stats
List:L3
Freq:1
C-Level:.9
Calculate
```



```
TInterval
(-4.22,37.554)
x̄=16.66666667
Sx=25.39028686
n=6
```

Example TI9–7

1. Enter the data values into L_1 and L_2.
2. Move the cursor to the top of the L_3 column so that L_3 is highlighted.
3. Type **$L_1 - L_2$,** then press **ENTER.**
4. Press **STAT** and move the cursor to TESTS.
5. Press **8** for TInterval.
6. Move the cursor to Data and press **ENTER.**
7. Type in the appropriate values, using **L_3** for the list.
8. Move the cursor to Calculate and press **ENTER.**

EXCEL Step by Step

Testing the Difference Between Two Means: Dependent Samples

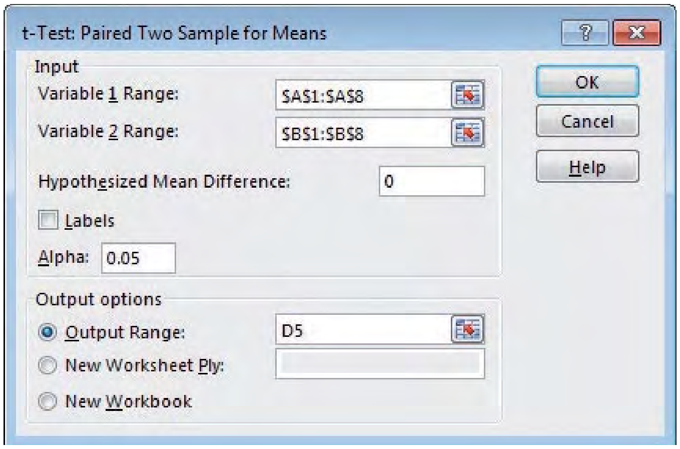
Example XL9–3

Test the claim that there is no difference between population means based on these sample paired data. Use $\alpha = 0.05$.

Set A	33	35	28	29	32	34	30	34
Set B	27	29	36	34	30	29	28	24

1. Enter the 8-number data set A into column A.
2. Enter the 8-number data set B into column B.
3. Select the Data tab from the toolbar. Then select Data Analysis.
4. In the Data Analysis box, under Analysis Tools select *t*-test: Paired Two Sample for Means, and click [OK].

5. In Input, type in the Variable 1 Range: **A1:A8** and the Variable 2 Range: **B1:B8.**
6. Type **0** for the Hypothesized Mean Difference.
7. Type **0.05** for Alpha.
8. In Output options, type **D5** for the Output Range, then click [OK].



t-Test: Paired Two Sample for Means

Input

Variable 1 Range: $A$1:$A$8

Variable 2 Range: $B$1:$B$8

Hypothesized Mean Difference: 0

☐ Labels

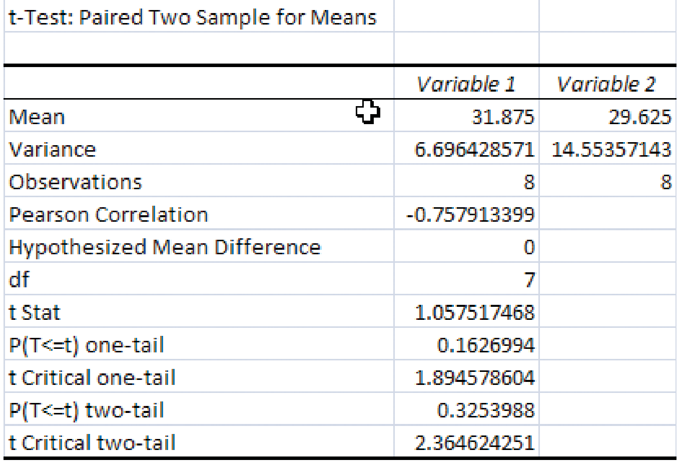
Alpha: 0.05

Output options

● Output Range: D5

○ New Worksheet Ply:

○ New Workbook



t-Test: Paired Two Sample for Means		
	Variable 1	*Variable 2*
Mean	31.875	29.625
Variance	6.696428571	14.55357143
Observations	8	8
Pearson Correlation	-0.757913399	
Hypothesized Mean Difference	0	
df	7	
t Stat	1.057517468	
P(T<=t) one-tail	0.1626994	
t Critical one-tail	1.894578604	
P(T<=t) two-tail	0.3253988	
t Critical two-tail	2.364624251	

Note: You may need to increase the column width to see all the results. To do this:

1. Highlight the columns D, E, and F.
2. Under the Home tab, select **Format>AutoFit** Column Width.

The output shows a *P*-value of 0.3253988 for the two-tailed case. This value is greater than the alpha level of 0.05, so we fail to reject the null hypothesis.

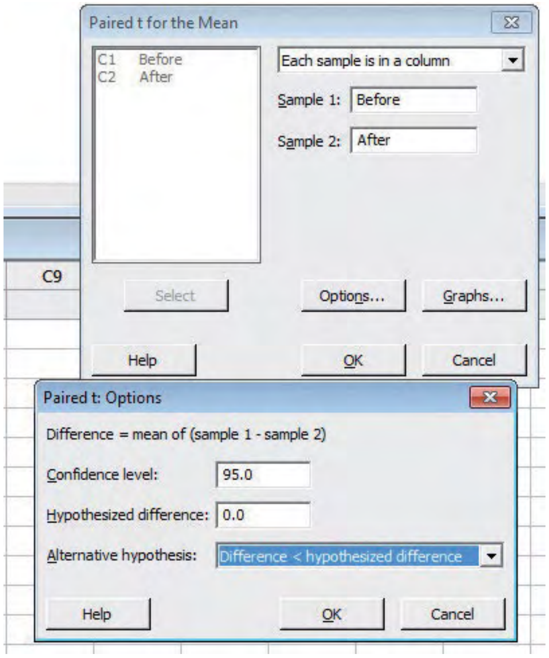
MINITAB
Step by Step

Test the Difference Between Two Means: Dependent Samples

A sports fitness trainer claims by taking a special vitamin, a weight lifter can increase his strength. Eight athletes are selected and given a test of strength, using the standard bench press. After 2 weeks of regular training, supplemented with the vitamin, they are tested again. Test the effectiveness of the vitamin regimen at $\alpha = 0.05$. Each value in these data represents the maximum number of pounds the athlete can bench-press. Assume that the variable is approximately normally distributed.

Athlete	1	2	3	4	5	6	7	8
Before (X_1)	210	230	182	205	262	253	219	216
After (X_2)	219	236	179	204	270	250	222	216

1. Enter the data into C1 and C2. Name the columns **Before** and **After.**
2. Select **Stat>Basic Statistics>Paired t.**
3. Select Each sample is in a column from the drop down menu.
4. Double click C1 Before for Sample 1.
5. Double click C2 After for Sample 2. The second sample will be subtracted from the first. The differences are not stored or displayed.
6. Click [Options].
7. Change Alternative hypothesis to Difference < hypothesized difference.
8. Click [OK] twice.


Paired t for the Mean

C1 Before
C2 After

Each sample is in a column

Sample 1: Before

Sample 2: After

Paired t: Options

Difference = mean of (sample 1 - sample 2)

Confidence level: 95.0

Hypothesized difference: 0.0

Alternative hypothesis: Difference < hypothesized difference

```
Paired T-Test and CI: Before, After
Paired T for Before – After
            N     Mean     StDev   SE Mean
Before      8   222.13     25.92      9.16
After       8   224.50     27.91      9.87
Difference  8    –2.38      4.84      1.71
95% upper bound for mean difference: 0.87
T-Test of mean difference = 0 (vs < 0) : T–Value = –1.39  P–Value = 0.104.
```

Since the *P*-value is 0.104, do not reject the null hypothesis. The sample difference of −2.38 in the strength measurement is not statistically significant.

9–4 Testing the Difference Between Proportions

OBJECTIVE 4

Test the difference between two proportions.

In Chapter 8, an inference about a single proportion was explained. In this section, testing the difference between two sample proportions will be explained.

The z test with some modifications can be used to test the equality of two proportions. For example, a researcher might ask, Is the proportion of men who exercise regularly less than the proportion of women who exercise regularly? Is there a difference in the percentage of students who own a mobile device and the percentage of nonstudents who own one? Is there a difference in the proportion of college graduates who pay cash for purchases and the proportion of non-college graduates who pay cash?

Recall from Chapter 7 that the symbol $\hat{p}$ ("p hat") is the sample proportion used to estimate the population proportion, denoted by p. For example, if in a sample of 30 college students, 9 are on probation, then the sample proportion is $\hat{p} = \frac{9}{30}$, or 0.3. The population proportion p is the number of all students who are on probation, divided by the number of students who attend the college. The formula for the sample proportion $\hat{p}$ is

$$\hat{p} = \frac{X}{n}$$

where

X = number of units that possess the characteristic of interest
n = sample size

When you are testing the difference between two population proportions p_1 and p_2, the hypotheses can be stated thus, if no specific difference between the proportions is hypothesized.

$$H_0\colon p_1 = p_2 \qquad \text{or} \qquad H_0\colon p_1 - p_2 = 0$$
$$H_1\colon p_1 \neq p_2 \qquad\qquad\quad H_1\colon p_1 - p_2 \neq 0$$

Similar statements using < or > in the alternate hypothesis can be formed for one-tailed tests.

For two proportions, $\hat{p}_1 = X_1/n_1$ is used to estimate p_1 and $\hat{p}_2 = X_2/n_2$ is used to estimate p_2. The standard error of the difference is

$$\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\sigma^2_{p_1} + \sigma^2_{p_2}} = \sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}$$

where $\sigma^2_{p_1}$ and $\sigma^2_{p_2}$ are the variances of the proportions, $q_1 = 1 - p_1$, $q_2 = 1 - p_2$, and n_1 and n_2 are the respective sample sizes.

Since p_1 and p_2 are unknown, a weighted estimate of p can be computed by using the formula

$$\overline{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$$

and $\overline{q} = 1 - \overline{p}$. This weighted estimate is based on the hypothesis that $p_1 = p_2$. Hence, $\overline{p}$ is a better estimate than either $\hat{p}_1$ or $\hat{p}_2$, since it is a combined average using both $\hat{p}_1$ and $\hat{p}_2$.

Since $\hat{p}_1 = X_1/n_1$ and $\hat{p}_2 = X_2/n_2$, $\overline{p}$ can be simplified to

$$\overline{p} = \frac{X_1 + X_2}{n_1 + n_2}$$

Finally, the standard error of the difference in terms of the weighted estimate is

$$\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\overline{p}\,\overline{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

The formula for the test value is shown next.

Formula for the *z* Test Value for Comparing Two Proportions

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\overline{p}\,\overline{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where

$$\overline{p} = \frac{X_1 + X_2}{n_1 + n_2} \qquad \hat{p}_1 = \frac{X_1}{n_1}$$

$$\overline{q} = 1 - \overline{p} \qquad \hat{p}_2 = \frac{X_2}{n_2}$$

This formula follows the format

$$\text{Test value} = \frac{\text{(observed value)} - \text{(expected value)}}{\text{standard error}}$$

Before you can test the difference between two sample proportions, the following assumptions must be met.

Assumptions for the *z* Test for Two Proportions

1. The samples must be random samples.
2. The sample data are independent of one another.
3. For both samples $np \geq 5$ and $nq \geq 5$.

In this book, the assumptions will be stated in the exercises; however, when encountering statistics in other situations, you must check to see that these assumptions have been met before proceeding.

The hypothesis-testing procedure used here follows the five-step procedure presented previously except that $\hat{p}_1$, $\hat{p}_2$, $\overline{p}$, and $\overline{q}$ must be computed.

EXAMPLE 9–9 Vaccination Rates in Nursing Homes

In the nursing home study mentioned in the chapter-opening Statistics Today, the researchers found that 12 out of 34 randomly selected small nursing homes had a resident vaccination rate of less than 80%, while 17 out of 24 randomly selected large nursing homes had a vaccination rate of less than 80%. At $\alpha = 0.05$, test the claim that there is no difference in the proportions of the small and large nursing homes with a resident vaccination rate of less than 80%.

Source: Nancy Arden, Arnold S. Monto, and Suzanne E. Ohmit, "Vaccine Use and the Risk of Outbreaks in a Sample of Nursing Homes During an Influenza Epidemic," *American Journal of Public Health.*

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon p_1 = p_2 \text{ (claim)} \qquad \text{and} \qquad H_1\colon p_1 \neq p_2$$

Step 2 Find the critical values. Since $\alpha = 0.05$, the critical values are +1.96 and −1.96.

Step 3 Compute the test value. First compute $\hat{p}_1$, $\hat{p}_2$, $\bar{p}$, and $\bar{q}$. Then substitute in the formula.

Let $\hat{p}_1$ be the proportion of the small nursing homes with a vaccination rate of less than 80% and $\hat{p}_2$ be the proportion of the large nursing homes with a vaccination rate of less than 80%. Then

$$\hat{p}_1 = \frac{X_1}{n_1} = \frac{12}{34} = 0.35 \qquad \text{and} \qquad \hat{p}_2 = \frac{X_2}{n_2} = \frac{17}{24} = 0.71$$

$$\bar{p} = \frac{X_1 + X_2}{n_1 + n_2} = \frac{12 + 17}{34 + 24} = \frac{29}{58} = 0.5$$

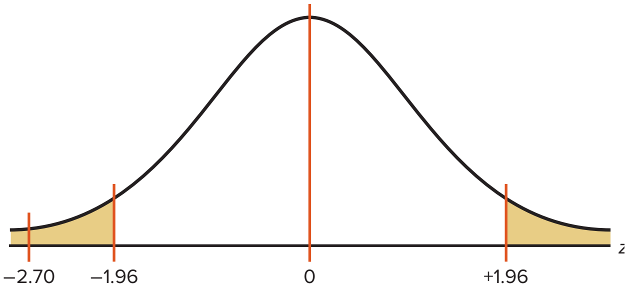
$$\bar{q} = 1 - \bar{p} = 1 - 0.5 = 0.5$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}\,\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{(0.35 - 0.71) - 0}{\sqrt{(0.5)(0.5)\left(\frac{1}{34} + \frac{1}{24}\right)}} = \frac{-0.36}{0.1333} = -2.70$$

Step 4 Make the decision. Reject the null hypothesis, since $-2.70 < -1.96$. See Figure 9–8.

FIGURE 9–8 Critical and Test Values for Example 9–9



Step 5 Summarize the results. There is enough evidence to reject the claim that there is no difference in the proportions of small and large nursing homes with a resident vaccination rate of less than 80%.

EXAMPLE 9–10 Criminal Arrests

A study found that in a random sample of 50 burglaries, 16% of the criminals were arrested. In a random sample of 50 car thefts, 12% of the criminals were arrested. At $\alpha = 0.10$, can it be concluded that the percentages of people who committed burglaries and were arrested was greater than the percentages of people who committed car thefts and were arrested.

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon p_1 = p_2 \qquad \text{and} \qquad H_1\colon p_1 > p_2 \text{ (claim)}$$

Step 2 Find the critical value, using Table E. At $\alpha = 0.10$ the critical value is 1.28.

Step 3 Compute the test value. Since percentages are given, you need to compare $\bar{p}$ and $\bar{q}$.

$$X_1 = \hat{p}_1 n_1 = 0.16(50) = 8$$

$$X_2 = \hat{p}_2 n_2 = 0.12(50) = 6$$

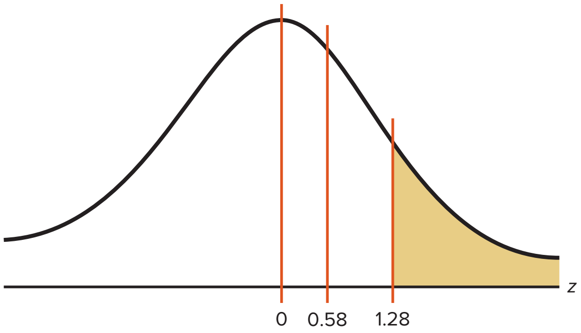
$$\bar{p} = \frac{X_1 + X_2}{n_1 + n_2} = \frac{8 + 6}{50 + 50} = \frac{14}{100} = 0.14$$

$$\bar{q} = 1 - \bar{p} = 1 - 0.14 = 0.86$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}\,\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.16 - 0.12) - 0}{\sqrt{(0.14)(0.86)\left(\frac{1}{50} + \frac{1}{50}\right)}} = \frac{0.04}{0.069} = 0.58$$

Step 4 Make the decision. Do not reject the null hypothesis since $0.58 < 1.28$. That is, 0.58 falls in the noncritical region. See Figure 9–9.

FIGURE 9–9 Critical and Test Value for Example 9–10



Step 5 Summarize the results. There is not enough evidence to support the claim that the percentage of people who are arrested for burglaries is greater than the percentage of people who are arrested who committed car thefts.

SPEAKING OF STATISTICS Is More Expensive Better?

An article in the *Journal of the American Medical Association* explained a study done on placebo pain pills. Researchers randomly assigned 82 healthy people to two groups. The individuals in the first group were given sugar pills, but they were told that the pills were a new, fast-acting opioid pain reliever similar to codeine and that they were listed at $2.50 each. The individuals in the other group received the same sugar pills but were told that the pills had been marked down to 10¢ each.

Each group received electrical shocks before and after taking the pills. They were then asked if the pills reduced the pain. Eighty-five percent of the group who were told that the pain pills cost $2.50 said that they were effective, while 61% of the group who received the supposedly discounted pills said that they were effective.

State possible null and alternative hypotheses for this study. What statistical test could be used in this study? What might be the conclusion of the study?

© Comstock/PictureQuest RF

The *P*-value for the difference of proportions can be found from Table E as shown in Section 9–1. In Example 9–10, the table value for 0.58 is 0.7190, and $1 - 0.7190 = 0.2810$. Hence, $0.2810 > 0.01$; thus the decision is to not reject the null hypothesis.

The sampling distribution of the difference of two proportions can be used to construct a confidence interval for the difference of two proportions. The formula for the confidence interval for the difference between two proportions is shown next.

Confidence Interval for the Difference Between Two Proportions

$$(\hat{p}_1 - \hat{p}_2) - z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} < p_1 - p_2 < (\hat{p}_1 - \hat{p}_2) + z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

Here, the confidence interval uses a standard deviation based on estimated values of the population proportions, but the hypothesis test uses a standard deviation based on the assumption that the two population proportions are equal. As a result, you may obtain different conclusions when using a confidence interval or a hypothesis test. So when testing for a difference of two proportions, you use the *z* test rather than the confidence interval.

EXAMPLE 9–11

Find the 95% confidence interval for the difference of proportions for the data in Example 9–9.

SOLUTION

$$\hat{p}_1 = \frac{12}{34} = 0.35 \qquad \hat{q}_1 = 0.65$$

$$\hat{p}_2 = \frac{17}{24} = 0.71 \qquad \hat{q}_2 = 0.29$$

Substitute in the formula.

$$(\hat{p}_1 - \hat{p}_2) - z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} < p_1 - p_2 < (\hat{p}_1 - \hat{p}_2) + z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

$$(0.35 - 0.71) - 1.96\sqrt{\frac{(0.35)(0.65)}{34} + \frac{(0.71)(0.29)}{24}} < p_1 - p_2 < (0.35 - 0.71) + 1.96\sqrt{\frac{(0.35)(0.65)}{34} + \frac{(0.71)(0.29)}{24}}$$

$$-0.36 - 0.242 < p_1 - p_2 < -0.36 + 0.242$$

$$-0.602 < p_1 - p_2 < -0.118$$

Applying the Concepts 9–4

Smoking and Education

You are researching the hypothesis that there is no difference in the percent of public school students who smoke and the percent of private school students who smoke. You find these results from a recent survey.

School	Percent who smoke
Public	32.3
Private	14.5

Based on these figures, answer the following questions.

1. What hypotheses would you use if you wanted to compare percentages of the public school students who smoke with the private school students who smoke?
2. What critical value(s) would you use?
3. What statistical test would you use to compare the two percentages?
4. What information would you need to complete the statistical test?
5. Suppose you found that 1000 randomly selected individuals in each group were surveyed. Could you perform the statistical test?
6. If so, complete the test and summarize the results.

See page 545 for the answers.

Exercises 9–4

1. Find the proportions $\hat{p}$ and $\hat{q}$ for each.

a. $n = 52, X = 32$
b. $n = 80, X = 66$
c. $n = 36, X = 12$
d. $n = 42, X = 7$
e. $n = 160, X = 50$

2. Find $\hat{p}$ and $\hat{q}$ for each.

a. $n = 36, X = 20$
b. $n = 50, X = 35$
c. $n = 64, X = 16$
d. $n = 200, X = 175$
e. $n = 148, X = 16$

3. Find each X, given $\hat{p}$.

a. $\hat{p} = 0.60, n = 240$
b. $\hat{p} = 0.20, n = 320$
c. $\hat{p} = 0.60, n = 520$
d. $\hat{p} = 0.80, n = 50$
e. $\hat{p} = 0.35, n = 200$

4. Find each X, given $\hat{p}$.

a. $\hat{p} = 0.24, n = 300$
b. $\hat{p} = 0.09, n = 200$
c. $\hat{p} = 88\%, n = 500$
d. $\hat{p} = 40\%, n = 480$
e. $\hat{p} = 32\%, n = 700$

5. Find $\hat{p}$ and $\hat{q}$ for each.

a. $X_1 = 25, n_1 = 75, X_2 = 40, n_2 = 90$
b. $X_1 = 9, n_1 = 15, X_2 = 7, n_2 = 20$
c. $X_1 = 3, n_1 = 20, X_2 = 5, n_2 = 40$
d. $X_1 = 21, n_1 = 50, X_2 = 32, n_2 = 50$
e. $X_1 = 20, n_1 = 150, X_2 = 30, n_2 = 50$

6. Find $\bar{p}$ and $\bar{q}$.

a. $X_1 = 6, n_1 = 15, X_2 = 9, n_2 = 15$
b. $X_1 = 21, n_1 = 100, X_2 = 43, n_2 = 150$
c. $X_1 = 20, n_1 = 80, X_2 = 65, n_2 = 120$
d. $X_1 = 15, n_1 = 50, X_2 = 3, n_2 = 12$
e. $X_1 = 24, n_1 = 40, X_2 = 18, n_2 = 36$

For Exercises 7 through 27, perform these steps.

a. State the hypotheses and identify the claim.
b. Find the critical value(s).
c. Compute the test value.
d. Make the decision.
e. Summarize the results.

Use the traditional method of hypothesis testing unless otherwise specified.

7. Lecture versus Computer-Assisted Instruction A survey found that 83% of the men questioned preferred computer-assisted instruction to lecture and 75% of the women preferred computer-assisted instruction to lecture. There were 100 randomly selected individuals in each sample. At $\alpha = 0.05$, test the claim that there is no difference in the proportion of men and the proportion of women who favor computer-assisted instruction over lecture. Find the 95% confidence interval for the difference of the two proportions.

8. Leisure Time In a sample of 150 men, 132 said that they had less leisure time today than they had 10 years ago. In a random sample of 250 women, 240 women said that they had less leisure time than they had 10 years ago. At $\alpha = 0.10$, is there a difference in the proportions? Find the 90% confidence interval for the difference of the two proportions. Does the confidence interval contain 0? Give a reason why this information would be of interest to a researcher.

Source: Based on statistics from Market Directory.

9. Desire to Be Rich In a random sample of 80 Americans, 44 wished that they were rich. In a random sample of 90 Europeans, 41 wished that they were rich. At $\alpha = 0.01$, is there a difference in the proportions? Find the 99% confidence interval for the difference of the two proportions.

10. Animal Bites of Postal Workers In Cleveland, a random sample of 73 mail carriers showed that 10 had been bitten by an animal during one week. In Philadelphia, in a random sample of 80 mail carriers, 16 had received animal bites. Is there a significant difference in the proportions? Use $\alpha = 0.05$. Find the 95% confidence interval for the difference of the two proportions.

11. Pet Ownership A recent random survey of households found that 14 out of 50 householders had a cat and 21 out of 60 householders had a dog. At $\alpha = 0.05$, test the claim that fewer household owners have cats than household owners who have dogs as pets.

12. Seat Belt Use In a random sample of 200 men, 130 said they used seat belts. In a random sample of 300 women, 63 said they used seat belts. Test the claim that men are more safety-conscious than women, at $\alpha = 0.01$. Use the P-value method.

13. Victims of Violence A random survey of 80 women who were victims of violence found that 24 were attacked by relatives. A random survey of 50 men found that 6 were attacked by relatives. At $\alpha = 0.10$, can it be shown that the percentage of women who were attacked by relatives is greater than the percentage of men who were attacked by relatives?

14. Hypertension It has been found that 26% of men 20 years and older suffer from hypertension (high blood pressure) and 31.5% of women are hypertensive. A random sample of 150 of each gender was selected from recent hospital records, and the following results were obtained. Can you conclude that a higher percentage of women have high blood pressure? Use $\alpha = 0.05$.

Men	43 patients had high blood pressure
Women	52 patients had high blood pressure

Source: www.nchs.gov

15. Commuters A recent random survey of 100 individuals in Michigan found that 80 drove to work alone. A similar survey of 120 commuters in

New York found that 62 drivers drove alone to work. Find the 95% confidence interval for the difference in proportions.

16. **Smoking Survey** National statistics show that 23% of men smoke and 18.5% of women smoke. A random sample of 180 men indicated that 50 were smokers, and a random sample of 150 women surveyed indicated that 39 smoked. Construct a 98% confidence interval for the true difference in proportions of male and female smokers. Comment on your interval—does it support the claim that there is a difference?

Source: www.nchs.gov

17. **Senior Workers** It seems that people are choosing or finding it necessary to work later in life. Random samples of 200 men and 200 women age 65 or older were selected, and 80 men and 59 women were found to be working. At $\alpha = 0.01$, can it be concluded that the proportions are different?

Source: Based on www.census.gov

18. **Airlines On-Time Arrivals** The percentages of on-time arrivals for major U.S. airlines range from 68.6 to 91.1. Two regional airlines were surveyed with the following results. At $\alpha = 0.01$, is there a difference in proportions?

	Airline A	Airline B
No. of flights	300	250
No. of on-time flights	213	185

Source: New York Times Almanac.

19. **College Education** The percentages of adults 25 years of age and older who have completed 4 or more years of college are 23.6% for females and 27.8% for males. A random sample of women and men who were 25 years old or older was surveyed with these results. Estimate the true difference in proportions with 95% confidence, and compare your interval with the *Almanac* statistics.

	Women	Men
Sample size	350	400
No. who completed 4 or more years	100	115

Source: New York Times Almanac.

20. **Married People** In a specific year 53.7% of men in the United States were married and 50.3% of women were married. Two independent random samples of 300 men and 300 women found that 178 men and 139 women were married (not to each other). At the 0.05 level of significance, can it be concluded that the proportion of men who were married is greater than the proportion of women who were married?

Source: New York Times Almanac.

21. **Undergraduate Financial Aid** A study is conducted to determine if the percent of women who receive financial aid in undergraduate school is different from the percent of men who receive financial aid in undergraduate school. A random sample of undergraduates revealed these results. At $\alpha = 0.01$, is there significant evidence to reject the null hypothesis?

	Women	Men
Sample size	250	300
Number receiving aid	200	180

Source: U.S. Department of Education, National Center for Education Statistics.

22. **High School Graduation Rates** The overall U.S. public high school graduation rate is 73.4%. For Pennsylvania it is 83.5% and for Idaho 80.5%—a difference of 3%. Random samples of 1200 students from each state indicated that 980 graduated in Pennsylvania and 940 graduated in Idaho. At the 0.05 level of significance, can it be concluded that there is a difference in the proportions of graduating students between the states?

Source: World Almanac.

23. **Interview Errors** It has been found that many first-time interviewees commit errors that could very well affect the outcome of the interview. An astounding 77% are guilty of using their cell phones or texting during the interview! A researcher wanted to see if the proportion of male offenders differed from the proportion of female ones. Out of 120 males, 72 used their cell phone and 80 of 150 females did so. At the 0.01 level of significance is there a difference?

Source: Careerbuilder.com

24. **Medical Supply Sales** According to the U.S. Bureau of Labor Statistics, approximately equal numbers of men and women are engaged in sales and related occupations. Although that may be true for total numbers, perhaps the proportions differ by industry. A random sample of 200 salespersons from the industrial sector indicated that 114 were men, and in the medical supply sector, 80 of 200 were men. At the 0.05 level of significance, can we conclude that the proportion of men in industrial sales differs from the proportion of men in medical supply sales?

25. **Coupon Use** In today's economy, everyone has become savings savvy. It is still believed, though, that a higher percentage of women than men clip coupons. A random survey of 180 female shoppers indicated that 132 clipped coupons while 56 out of 100 men did so. At $\alpha = 0.01$, is there sufficient evidence that the proportion of couponing women is higher than the proportion of couponing men? Use the *P*-value method.

26. **Never Married People** The percentage of males 18 years and older who have never married is 30.4. For females the percentage is 23.6. Looking at the records in a particular populous county, a random sample of 250 men showed that 78 had never married and 58 of 200 women had never married. At the 0.05 level of significance, is the proportion of men greater than the proportion of women? Use the *P*-value method.

27. Bullying Bullying is a problem at any age but especially for students aged 12 to 18. A study showed that 7.2% of all students in this age bracket reported being bullied at school during the past six months with 6th grade having the highest incidence at 13.9% and 12th grade the lowest at 2.2%. To see if there is a difference between public and private schools, 200 students were randomly selected from each. At the 0.05 level of significance, can a difference be concluded?

	Private	**Public**
Sample size	200	200
No. bullied	13	16

Source: www.nces.ed.gov

Extending the Concepts

28. If there is a significant difference between p_1 and p_2 and between p_2 and p_3, can you conclude that there is a significant difference between p_1 and p_3?

Technology Step by Step

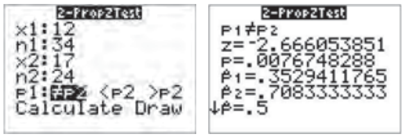
TI-84 Plus Step by Step

Hypothesis Test for the Difference Between Two Proportions

Example TI9–8

1. Press **STAT** and move the cursor to TESTS.
2. Press **6** for 2-PropZTEST.
3. Type in the appropriate values.
4. Move the cursor to the appropriate alternative hypothesis and press **ENTER.**
5. Move the cursor to Calculate and press **ENTER.**

This refers to Example 9–9 in the text.



```
2-PropZTest
x1:12
n1:34
x2:17
n2:24
p1:≠p2 <p2 >p2
Calculate Draw
```

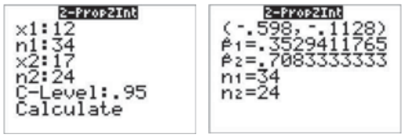
```
2-PropZTest
p1≠p2
z=⁻2.666053851
p=.0076748288
p̂1=.3529411765
p̂2=.7083333333
↓p̂=.5
```

Confidence Interval for the Difference Between Two Proportions

Example TI9–9

1. Press **STAT** and move the cursor to TESTS.
2. Press **B (ALPHA APPS)** for 2-PropZInt.
3. Type in the appropriate values.
4. Move the cursor to Calculate and press **ENTER.**

This refers to Example 9–11 in the text.



```
2-PropZInt
x1:12
n1:34
x2:17
n2:24
C-Level:.95
Calculate
```

```
2-PropZInt
(-.598,-.1128)
p̂1=.3529411765
p̂2=.7083333333
n1=34
n2=24
```

EXCEL Step by Step

Testing the Difference Between Two Proportions

Excel does not have a procedure to test the difference between two population proportions. However, you may conduct this test using the MegaStat Add-in available in your online resources. If you have not installed this add-in, do so, following the instructions from the Chapter 1 Excel Step by Step.

We will use the summary information from Example 9–9.

1. From the toolbar, select Add-Ins, **MegaStat>Hypothesis Tests>Compare Two Independent Proportions.** *Note:* You may need to open MegaStat from the MegaStat.xls file on your computer's hard drive.
2. Under Group 1, type **12** for *p* and **34** for *n*. Under Group 2, type **17** for *p* and **24** for *n*. MegaStat automatically changes *p* to *X* unless a decimal value less than 1 is typed in for these.
3. Type **0** for the Hypothesized difference, select the not equal Alternative, and click [OK].

Hypothesis Test for Two Independent Proportions

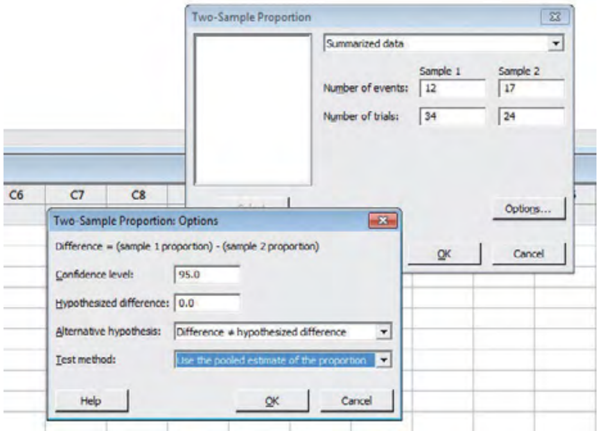
```
        p1          p2                                pc
    0.3529      0.7083                               0.5    p (as decimal)
     12/34       17/24                             29/58    p (as fraction)
       12.         17.                               29.    X
        34          24                                58    n
               −0.3554   Difference
                    0.   Hypothesized difference
                0.1333   Standard error
                 −2.67   z
                0.0077   P-value (two-tailed)
```

MINITAB
Step by Step

Test the Difference Between Two Proportions

For Example 9–9, test for a difference in the resident vaccination rates between small and large nursing homes.

1. This test does not require data. It doesn't matter what is in the worksheet.
2. Select Summarized data from the drop down menu.
3. Click the button for Summarized data.
4. Press **TAB** to move cursor to the Sample 1 box for Number of events.
   a) Enter **12, TAB,** then enter **34.**
   b) Press **TAB** or click in the Sample 2 text box for Number of events.
   c) Enter **17, TAB,** then enter **24.**
5. Click on [Options]. The Confidence level should be 95%, and the Hypothesized difference should be 0.
   a) For the Alternative hypothesis, select Difference ≠ hypothesized difference.
   b) For the Test method, select Use the pooled estimate of the proportion.
6. Click [OK] twice. The results are shown in the session window.



Two-Sample Proportion

Summarized data

	Sample 1	Sample 2
Number of events:	12	17
Number of trials:	34	24

Two-Sample Proportion: Options

Difference = (sample 1 proportion) - (sample 2 proportion)

Confidence level: 95.0

Hypothesized difference: 0.0

Alternative hypothesis: Difference ≠ hypothesized difference

Test method: Use the pooled estimate of the proportion

```
Test and CI for Two Proportions
Sample   X   N   Sample p
1       12  34   0.352941
2       17  24   0.708333

Difference = p (1) – p (2)
Estimate for difference:  −0.355392
95% CI for difference:  (−0.598025, −0.112759)
Test for difference = 0 (vs ≠ 0):  Z = −2.67  P–Value = 0.008
```

The *P*-value of the test is 0.008. Reject the null hypothesis. The difference is statistically significant. Of all small nursing homes 35%, compared to 71% of all large nursing homes, have an immunization rate of less than 80%. We can't tell why, only that there is a difference. ▬

9–5 Testing the Difference Between Two Variances

OBJECTIVE ⑤

Test the difference between two variances or standard deviations.

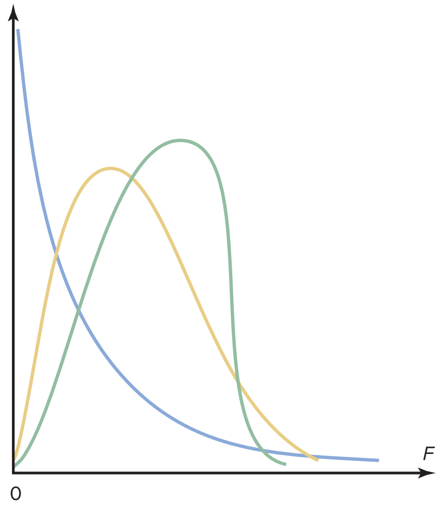
In addition to comparing two means, statisticians are interested in comparing two variances or standard deviations. For example, is the variation in the temperatures for a certain month for two cities different?

In another situation, a researcher may be interested in comparing the variance of the cholesterol of men with the variance of the cholesterol of women. For the comparison of two variances or standard deviations, an ***F* test** is used. The *F* test should not be confused with the chi-square test, which compares a single sample variance to a specific population variance, as shown in Chapter 8.

Figure 9–10 shows the shapes of several curves for the *F* distribution.

FIGURE 9–10

The *F* Family of Curves



If two independent samples are selected from two normally distributed populations in which the population variances are equal ($\sigma_1^2 = \sigma_2^2$) and if the sample variances s_1^2 and s_2^2 are compared as $\frac{s_1^2}{s_2^2}$, the sampling distribution of the variances is called the ***F* distribution.**

Characteristics of the *F* Distribution

1. The values of *F* cannot be negative, because variances are always positive or zero.
2. The distribution is positively skewed.
3. The mean value of *F* is approximately equal to 1.
4. The *F* distribution is a family of curves based on the degrees of freedom of the variance of the numerator and the degrees of freedom of the variance of the denominator.

Formula for the *F* Test

$$F = \frac{s_1^2}{s_2^2}$$

where the larger of the two variances is placed in the numerator regardless of the subscripts. (See note on page 534.)

The *F* test has two values for the degrees of freedom: that of the numerator, $n_1 - 1$, and that of the denominator, $n_2 - 1$, where n_1 is the sample size from which the larger variance was obtained.

When you are finding the *F* test value, *the larger of the variances is placed in the numerator of the F formula;* this is not necessarily the variance of the larger of the two sample sizes.

Table H in Appendix A gives the *F* critical values for $\alpha = 0.005$, 0.01, 0.025, 0.05, and 0.10 (each α value involves a separate table in Table H). These are one-tailed values; if a two-tailed test is being conducted, then the $\alpha/2$ value must be used. For example, if a two-tailed test with $\alpha = 0.05$ is being conducted, then the $0.05/2 = 0.025$ table of Table H should be used.

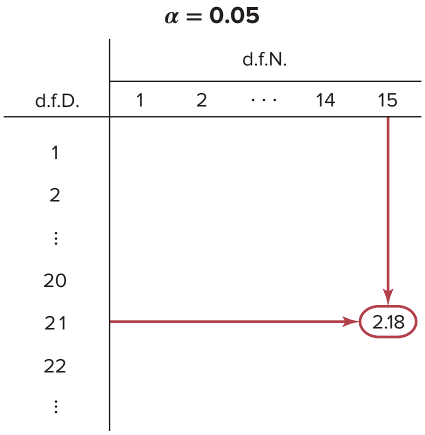
EXAMPLE 9–12

Find the critical value for a right-tailed *F* test when $\alpha = 0.05$, the degrees of freedom for the numerator (abbreviated d.f.N.) are 15, and the degrees of freedom for the denominator (d.f.D.) are 21.

SOLUTION

Since this test is right-tailed with $\alpha = 0.05$, use the 0.05 table. The d.f.N. is listed across the top, and the d.f.D. is listed in the left column. The critical value is found where the row and column intersect in the table. In this case, it is 2.18. See Figure 9–11.

FIGURE 9–11 Finding the Critical Value in Table H for Example 9–12



As noted previously, when the F test is used, the larger variance is always placed in the numerator of the formula. When you are conducting a two-tailed test, α is split; and even though there are two values, only the right tail is used. The reason is that the F test value is always greater than or equal to 1.

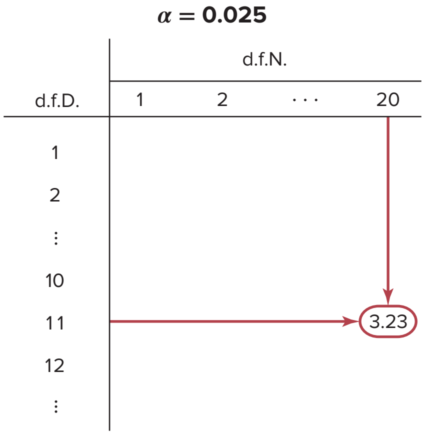
EXAMPLE 9–13

Find the critical value for a two-tailed F test with $\alpha = 0.05$ when the sample size from which the variance for the numerator was obtained was 21 and the sample size from which the variance for the denominator was obtained was 12.

SOLUTION

Since this is a two-tailed test with $\alpha = 0.05$, the $0.05/2 = 0.025$ table must be used. Here, d.f.N. $= 21 - 1 = 20$, and d.f.D. $= 12 - 1 = 11$; hence, the critical value is 3.23. See Figure 9–12.

FIGURE 9–12 Finding the Critical Value in Table H for Example 9–13



If the exact degrees of freedom are not specified in Table H, the closest smaller value should be used. For example, if $\alpha = 0.05$ (right-tailed test), d.f.N. = 18, and d.f.D. = 20, use the column d.f.N. = 15 and the row d.f.D. = 20 to get $F = 2.20$. Using the smaller value is the more conservative approach.

When you are testing the equality of two variances, these hypotheses are used:

Right-tailed	Left-tailed	Two-tailed
$H_0\colon \sigma_1^2 = \sigma_2^2$ $H_1\colon \sigma_1^2 > \sigma_2^2$	$H_0\colon \sigma_1^2 = \sigma_2^2$ $H_1\colon \sigma_1^2 < \sigma_2^2$	$H_0\colon \sigma_1^2 = \sigma_2^2$ $H_1\colon \sigma_1^2 \neq \sigma_2^2$

There are four key points to keep in mind when you are using the F test.

Notes for the Use of the *F* Test

1. The larger variance should always be placed in the numerator of the formula regardless of the subscripts. (See note on page 534.)

$$F = \frac{s_1^2}{s_2^2}$$

2. For a two-tailed test, the α value must be divided by 2 and the critical value placed on the right side of the F curve.
3. If the standard deviations instead of the variances are given in the problem, they must be squared for the formula for the F test.
4. When the degrees of freedom cannot be found in Table H, the closest value on the smaller side should be used.

Unusual Stat

Of all U.S. births, 2% are twins.

Before you can use the testing method to determine the difference between two variances, the following assumptions must be met.

Assumptions for Testing the Difference Between Two Variances

1. The samples must be random samples.
2. The populations from which the samples were obtained must be normally distributed. (*Note:* The test should not be used when the distributions depart from normality.)
3. The samples must be independent of one another.

In this book, the assumptions will be stated in the exercises; however, when encountering statistics in other situations, you must check to see that these assumptions have been met before proceeding.

Remember also that in tests of hypotheses using the traditional method, these five steps should be taken:

Step 1 State the hypotheses and identify the claim.

Step 2 Find the critical value.

Step 3 Compute the test value.

Step 4 Make the decision.

Step 5 Summarize the results.

This procedure is not robust, so minor departures from normality will affect the results of the test. So this test should not be used when the distributions depart from normality because standard deviations are not a good measure of the spread in nonsymmetrical distributions. The reason is that the standard deviation is not resistant to outliers or extreme values. These values increase the value of the standard deviation when the distribution is skewed.

EXAMPLE 9–14 Heart Rates of Smokers

A medical researcher wishes to see whether the variance of the heart rates (in beats per minute) of smokers is different from the variance of heart rates of people who do not smoke. Two samples are selected, and the data are shown. Using $\alpha = 0.05$, is there enough evidence to support the claim? Assume the variable is normally distributed.

Smokers	Nonsmokers
$n_1 = 26$ $s_1^2 = 36$	$n_2 = 18$ $s_2^2 = 10$

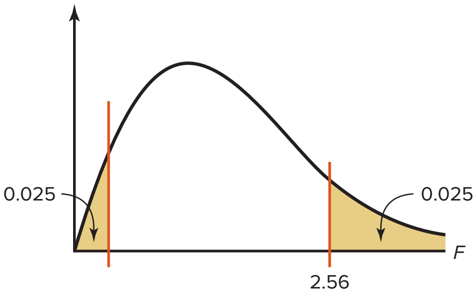
SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon \sigma_1^2 = \sigma_2^2 \qquad \text{and} \qquad H_1\colon \sigma_1^2 \neq \sigma_2^2 \text{ (claim)}$$

Step 2 Find the critical value. Use the 0.025 table in Table H since $\alpha = 0.05$ and this is a two-tailed test. Here, d.f.N. = 26 − 1 = 25, and d.f.D. = 18 − 1 = 17. The critical value is 2.56 (d.f.N. = 24 was used). See Figure 9–13.

FIGURE 9–13 Critical Value for Example 9–14



Step 3 Compute the test value.

$$F = \frac{s_1^2}{s_2^2} = \frac{36}{10} = 3.6$$

Step 4 Make the decision. Reject the null hypothesis, since 3.6 > 2.56.

Step 5 Summarize the results. There is enough evidence to support the claim that the variance of the heart rates of smokers and nonsmokers is different.

EXAMPLE 9–15 Grade Point Averages

A researcher selected a random sample of 10 psychology graduates and found the standard deviation of their grade point average was 0.72. Then she selected a random sample of 14 engineering students and found that the standard deviation of their grade point average was 0.51. At $\alpha = 0.01$, can we conclude that the variance of the grade point averages of the psychology graduates is greater than the variance of the grade point averages of the engineering graduates?

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon \sigma_1^2 = \sigma_2^2 \qquad \text{and} \qquad H_1\colon \sigma_1^2 > \sigma_2^2 \text{ (claim)}$$

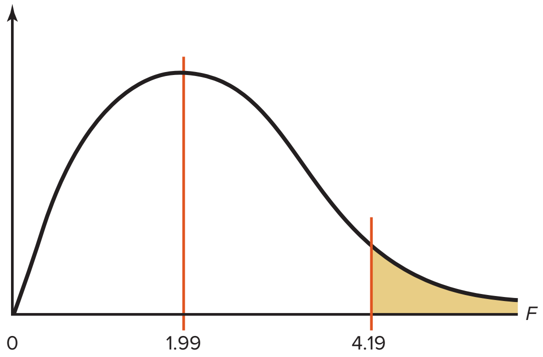
Step 2 Find the critical value. Hence, d.f.N. = 10 − 1 = 9 and d.f.D = 14 − 1 = 13. From Table H at $\alpha = 0.01$ the critical value is 4.19.

Step 3 Compute the test value.

$$F = \frac{s_1^2}{s_2^2} = \frac{0.72^2}{0.51^2} = 1.99$$

Step 4 Make the decision. Do not reject the null hypothesis since 1.99 < 4.19. That is, 1.99 does not fall in the critical region. See Figure 9–14.

FIGURE 9–14 Critical and Test Value for Example 9–15



Step 5 Summarize the results. There is not enough evidence to support the claim that the variance in the grade point average of psychology graduates is greater than the variance in the grade point average of the engineering graduates.

Finding *P*-values for the *F* test statistic is somewhat more complicated since it requires looking through all the *F* tables (Table H in Appendix A) using the specific d.f.N. and d.f.D. values. For example, suppose that a certain test has $F = 3.58$, d.f.N. = 5, and d.f.D. = 10. To find the *P*-value interval for $F = 3.58$, you must first find the corresponding *F* values for d.f.N. = 5 and d.f.D. = 10 for α equal to 0.005, 0.01, 0.025, 0.05, and 0.10 in Table H. Then make a table as shown.

α	0.10	0.05	0.025	0.01	0.005
F	2.52	3.33	4.24	5.64	6.87

Now locate the two *F* values that the test value 3.58 falls between. In this case, 3.58 falls between 3.33 and 4.24, corresponding to 0.05 and 0.025. Hence, the *P*-value for a right-tailed test for $F = 3.58$ falls between 0.025 and 0.05 (that is, $0.025 < P\text{-value} < 0.05$). For a right-tailed test, then, you would reject the null hypothesis at $\alpha = 0.05$, but not at $\alpha = 0.01$. The *P*-value obtained from a calculator is 0.0408. Remember that for a two-tailed test the values found in Table H for α must be doubled. In this case, $0.05 < P\text{-value} < 0.10$ for $F = 3.58$. Once again, if the *P*-value is less than α, we reject the null hypothesis.

Once you understand the concept, you can dispense with making a table as shown and find the *P*-value directly from Table H.

EXAMPLE 9–16 Airport Passengers

The CEO of an airport hypothesizes that the variance in the number of passengers for American airports is greater than the variance in the number of passengers for foreign airports. At $\alpha = 0.10$, is there enough evidence to support the hypothesis? The data in millions of passengers per year are shown for selected airports. Use the *P*-value method. Assume the variable is normally distributed and the samples are random and independent.

American airports		Foreign airports	
36.8	73.5	60.7	51.2
72.4	61.2	42.7	38.6
60.5	40.1		

Source: Airports Council International.

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0: \sigma_1^2 = \sigma_2^2 \qquad \text{and} \qquad H_1: \sigma_1^2 > \sigma_2^2 \text{ (claim)}$$

Step 2 Compute the test value. Using the formula in Chapter 3 or a calculator, find the variance for each group.

$$s_1^2 = 246.38 \qquad \text{and} \qquad s_2^2 = 95.87$$

Substitute in the formula and solve.

$$F = \frac{s_1^2}{s_2^2} = \frac{246.38}{95.87} = 2.57$$

Step 3 Find the *P*-value in Table H, using d.f.N. = 6 − 1 = 5 and d.f.D. = 4 − 1 = 3.

α	0.10	0.05	0.025	0.01	0.005
F	5.31	9.01	14.88	28.24	45.39

Since 2.57 is less than 5.31, the *P*-value is greater than 0.10. (The *P*-value obtained from a calculator is 0.234.)

Step 4 Make the decision. The decision is to not reject the null hypothesis since *P*-value > 0.10.

Step 5 Summarize the results. There is not enough evidence to support the claim that the variance in the number of passengers for American airports is greater than the variance in the number of passengers for foreign airports.

Note: It is not absolutely necessary to place the larger variance in the numerator when you are performing the *F* test. Critical values for left-tailed hypotheses tests can be found by interchanging the degrees of freedom and taking the reciprocal of the value found in Table H.

Also, you should use caution when performing the *F* test since the data can run contrary to the hypotheses on rare occasions. For example, if the hypotheses are H_0: $\sigma_1^2 \leq \sigma_2^2$ (written H_0: $\sigma_1^2 = \sigma_2^2$) and H_1: $\sigma_1^2 > \sigma_2^2$, but if $s_1^2 < s_2^2$, then the *F* test should not be performed and you would not reject the null hypothesis.

Applying the Concepts 9–5

Automatic Transmissions

Assume the following data values are from a 2016 Auto Guide. The guide compared various parameters of U.S.- and foreign-made cars. This report centers on the price of an optional automatic transmission. Which country has the greater variability in the price of automatic transmissions? Answer the following questions.

Germany (2016)		U.S. Cars (2016)	
BMW 6 Series	$77,300	Ford Mustang	$47,795
Audi TT	$46,400	Chevrolet Corvette	$55,400
Porsche Boxster	$82,100	Dodge Challenger	$62,495
BMW 2 Series	$50,750	Dodge Viper	$87,895

1. What is the null hypothesis?
2. What test statistic is used to test for any significant differences in the variances?

3. Is there a significant difference in the variability in the prices between the German cars and the U.S. cars?
4. What effect does a small sample size have on the standard deviations?
5. What degrees of freedom are used for the statistical test?
6. Could two sets of data have significantly different variances without having significantly different means?

See page 545 for the answers.

Exercises 9–5

1. When one is computing the F test value, what condition is placed on the variance that is in the numerator?

2. Why is the critical region always on the right side in the use of the F test?

3. What are the two different degrees of freedom associated with the F distribution?

4. What are the characteristics of the F distribution?

5. Using Table H, find the critical value for each.

a. Sample 1: $s_1^2 = 140$, $n_1 = 25$
Sample 2: $s_2^2 = 125$, $n_2 = 14$
Two-tailed $\alpha = 0.05$

b. Sample 1: $s_1^2 = 43$, $n_1 = 12$
Sample 2: $s_2^2 = 56$, $n_2 = 16$
Right-tailed $\alpha = 0.10$

c. Sample 1: $s_1^2 = 516$, $n_1 = 21$
Sample 2: $s_2^2 = 472$, $n_2 = 18$
Right-tailed $\alpha = 0.01$

6. Using Table H, find the critical value for each.

a. Sample 1: $s_1^2 = 27.3$, $n_1 = 5$
Sample 2: $s_2^2 = 38.6$, $n_2 = 9$
Right-tailed, $\alpha = 0.01$

b. Sample 1: $s_1^2 = 164$, $n_1 = 21$
Sample 2: $s_2^2 = 53$, $n_2 = 17$
Two-tailed, $\alpha = 0.10$

c. Sample 1: $s_1^2 = 92.8$, $n_1 = 11$
Sample 2: $s_2^2 = 43.6$, $n_2 = 11$
Right-tailed, $\alpha = 0.05$

7. Using Table H, find the P-value interval for each F test value.

a. $F = 2.97$, d.f.N. = 9, d.f.D. = 14, right-tailed
b. $F = 3.32$, d.f.N. = 6, d.f.D. = 12, two-tailed
c. $F = 2.28$, d.f.N. = 12, d.f.D. = 20, right-tailed
d. $F = 3.51$, d.f.N. = 12, d.f.D. = 21, right-tailed

8. Using Table H, find the P-value interval for each F test value.

a. $F = 4.07$, d.f.N. = 6, d.f.D. = 10, two-tailed
b. $F = 1.65$, d.f.N. = 19, d.f.D. = 28, right-tailed
c. $F = 1.77$, d.f.N. = 28, d.f.D. = 28, right-tailed
d. $F = 7.29$, d.f.N. = 5, d.f.D. = 8, two-tailed

For Exercises 9 through 24, perform the following steps. Assume that all variables are normally distributed.

a. State the hypotheses and identify the claim.
b. Find the critical value.
c. Compute the test value.
d. Make the decision.
e. Summarize the results.

Use the traditional method of hypothesis testing unless otherwise specified.

9. Wolf Pack Pups Does the variance in the number of pups per pack differ between Montana and Idaho wolf packs? Random samples of packs were selected for each area, and the numbers of pups per pack were recorded. At the 0.05 level of significance, can a difference in variances be concluded?

Montana wolf packs	4	3	5	6	1	2	8	2
	3	1	7	6	5			
Idaho wolf packs	2	4	5	4	2	4	6	3
	1	4	2	1				

Source: www.fws.gov

10. Noise Levels in Hospitals In a hospital study, it was found that the standard deviation of the sound levels from 20 randomly selected areas designated as "casualty doors" was 4.1 dBA and the standard deviation of 24 randomly selected areas designated as operating theaters was 7.5 dBA. At $\alpha = 0.05$, can you substantiate the claim that there is a difference in the standard deviations?

Source: M. Bayo, A. Garcia, and A. Garcia, "Noise Levels in an Urban Hospital and Workers' Subjective Responses," *Archives of Environmental Health.*

11. Calories in Ice Cream The numbers of calories contained in $\frac{1}{2}$-cup servings of randomly selected flavors of ice cream from two national brands are listed. At the 0.05 level of significance, is there sufficient evidence to conclude that the variance in the number of calories differs between the two brands?

Brand A		**Brand B**	
330	300	280	310
310	350	300	370
270	380	250	300
310	300	290	310

Source: The Doctor's Pocket Calorie, Fat and Carbohydrate Counter.

12. Winter Temperatures A random sample of daily high temperatures in January and February is listed. At $\alpha = 0.05$, can it be concluded that there is a difference in variances in high temperature between the two months?

Jan.	31	31	38	24	24	42	22	43	35	42
Feb.	31	29	24	30	28	24	27	34	27	

13. Population and Area Cities were randomly selected from the list of the 50 largest cities in the United States (based on population). The areas of each in square miles are shown. Is there sufficient evidence to conclude that the variance in area is greater for eastern cities than for western cities at $\alpha = 0.05$? At $\alpha = 0.01$?

Eastern		**Western**	
Atlanta, GA	132	Albuquerque, NM	181
Columbus, OH	210	Denver, CO	155
Louisville, KY	385	Fresno, CA	104
New York, NY	303	Las Vegas, NV	113
Philadelphia, PA	135	Portland, OR	134
Washington, DC	61	Seattle, WA	84
Charlotte, NC	242		

Source: New York Times Almanac.

14. Carbohydrates in Candy The number of grams of carbohydrates contained in 1-ounce servings of randomly selected chocolate and nonchocolate candy is shown. Is there sufficient evidence to conclude that there is a difference between the variation in carbohydrate content for chocolate and nonchocolate candy? Use $\alpha = 0.10$.

Chocolate	29	25	17	36	41	25	32	29
	38	34	24	27	29			
Nonchocolate	41	41	37	29	30	38	39	10
	29	55	29					

Source: The Doctor's Pocket Calorie, Fat and Carbohydrate Counter.

15. Tuition Costs for Medical School The yearly tuition costs in dollars for random samples of medical schools that specialize in research and in primary care are listed. At $\alpha = 0.05$, can it be concluded that a difference between the variances of the two groups exists?

Research			**Primary care**		
30,897	34,280	31,943	26,068	21,044	30,897
34,294	31,275	29,590	34,208	20,877	29,691
20,618	20,500	29,310	33,783	33,065	35,000
21,274			27,297		

Source: U.S. News & World Report Best Graduate Schools.

16. County Size in Indiana and Iowa A researcher wishes to see if the variance of the areas in square miles for counties in Indiana is less than the variance of the areas for counties in Iowa. A random sample of counties is selected, and the data are shown. At $\alpha = 0.01$, can it be concluded that the variance of the areas for counties in Indiana is less than the variance of the areas for counties in Iowa?

Indiana				**Iowa**			
406	393	396	485	640	580	431	416
431	430	369	408	443	569	779	381
305	215	489	293	717	568	714	731
373	148	306	509	571	577	503	501
560	384	320	407	568	434	615	402

Source: The World Almanac and Book of Facts.

17. Heights of Tall Buildings Test the claim that the variance of heights of randomly selected tall buildings in Denver is equal to the variance in heights of randomly selected tall buildings in Nashville at $\alpha = 0.10$. The data are given in feet.

Denver			**Nashville**		
714	698	544	617	524	489
504	438	408	459	453	417
404			410	404	

Source: SkyscraperCenter.com

18. Reading Program Summer reading programs are very popular with children. At the Citizens Library, Team Ramona read an average of 23.2 books with a standard deviation of 6.1. There were 21 members on this team. Team Beezus read an average of 26.1 books with a standard deviation of 2.3. There were 23 members on this team. Did the variances of the two teams differ? Use $\alpha = 0.05$.

19. Weights of Running Shoes The weights in ounces of a random sample of running shoes for men and women are shown. Calculate the variances for each sample, and test the claim that the variances are equal at $\alpha = 0.05$. Use the P-value method.

Men			**Women**		
11.9	10.4	12.6	10.6	10.2	8.8
12.3	11.1	14.7	9.6	9.5	9.5
9.2	10.8	12.9	10.1	11.2	9.3
11.2	11.7	13.3	9.4	10.3	9.5
13.8	12.8	14.5	9.8	10.3	11.0

20. School Teachers' Salaries A researcher claims that the variation in the salaries of elementary school teachers is greater than the variation in the salaries of secondary school teachers. A random sample of the salaries of 30 elementary school teachers has a variance of 8324, and a random sample of the salaries of 30 secondary school teachers has a variance of 2862. At $\alpha = 0.05$, can the researcher conclude that the variation in the elementary school teachers' salaries is greater than the variation in the secondary school teachers' salaries? Use the *P*-value method.

21. Ages of Dogs The average age of pet dogs is 4.3 years. Two random samples of pet owners who own dogs are selected. Sample 1 of 13 dog owners was selected from owners who live in Miami. The standard deviation of the ages of the dogs in this sample is 1.3 years. Sample 2 of 8 dog owners was selected from dog owners who live in Boston. The standard deviation of these dogs was 0.7 year. At $\alpha = 0.05$, can it be concluded that there is a difference in the variances?

22. Daily Stock Prices Two portfolios were randomly assembled from the New York Stock Exchange, and the daily stock prices are shown. At the 0.05, level of significance, can it be concluded that a difference in variance in price exists between the two portfolios?

23. Test Scores An instructor who taught an online statistics course and a classroom course feels that the variance of the final exam scores for the students who took the online course is greater than the variance of the final exam scores of the students who took the classroom final exam. The following data were obtained. At $\alpha = 0.05$ is there enough evidence to support the claim?

Online Course	Classroom Course
$s_1 = 3.2$	$s_2 = 2.8$
$n_1 = 11$	$n_2 = 16$

24. Museum Attendance A metropolitan children's museum open year-round wants to see if the variance in daily attendance differs between the summer and winter months. Random samples of 30 days each were selected and showed that in the winter months, the sample mean daily attendance was 300 with a standard deviation of 52, and the sample mean daily attendance for the summer months was 280 with a standard deviation of 65. At $\alpha = 0.05$, can we conclude a difference in variances?

Portfolio A	36.44	44.21	12.21	59.60	55.44	39.42	51.29	48.68	41.59	19.49
Portfolio B	32.69	47.25	49.35	36.17	63.04	17.74	4.23	34.98	37.02	31.48

Source: Washington Observer-Reporter.

Technology Step by Step

TI-84 Plus Step by Step

Hypothesis Test for the Difference Between Two Variances (Data)

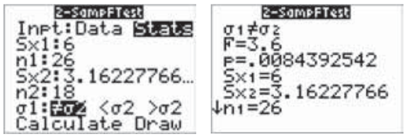
1. Enter the data values into L_1 and L_2.
2. Press **STAT** and move the cursor to TESTS.
3. Press **E (ALPHA SIN)** for 2-SampFTest.
4. Move the cursor to Data and press **ENTER.**
5. Type in the appropriate values.
6. Move the cursor to the appropriate Alternative hypothesis and press **ENTER.**
7. Move the cursor to Calculate and press **ENTER.**

Hypothesis Test for the Difference Between Two Variances (Statistics)

Example TI9–10

1. Press **STAT** and move the cursor to TESTS.
2. Press **E (ALPHA SIN)** for 2-SampFTest.
3. Move the cursor to Stats and press **ENTER.**
4. Type in the appropriate values.
5. Move the cursor to the appropriate Alternative hypothesis and press **ENTER.**
6. Move the cursor to Calculate and press **ENTER.**

This refers to Example 9–14 in the text.



```
2-SampFTest
Inpt:Data Stats
Sx1:6
n1:26
Sx2:3.16227766...
n2:18
σ1:≠σ2 <σ2 >σ2
Calculate Draw
```

```
2-SampFTest
σ1≠σ2
F=3.6
p=.0084392542
Sx1=6
Sx2=3.16227766
↓n1=26
```

EXCEL
Step by Step

F Test for the Difference Between Two Variances

Excel has a two-sample *F* test included in the Data Analysis Add-in. To perform an *F* test for the difference between the variances of two populations, given two independent samples, do this:

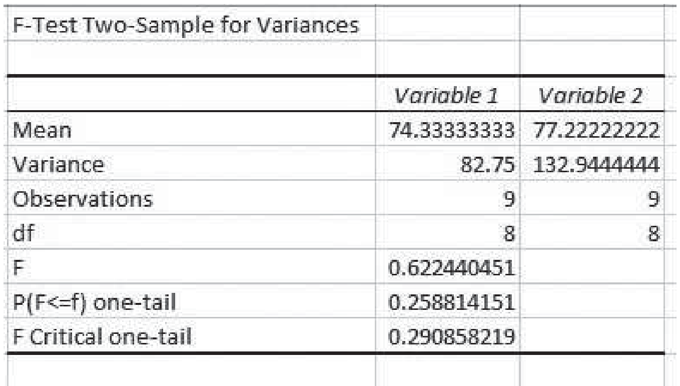
1. Enter the first sample data set into column A.
2. Enter the second sample data set into column B.
3. Select the Data tab from the toolbar. Then select Data Analysis.
4. In the Analysis Tools box, select F-test Two-sample for Variances.
5. Type the ranges for the data in columns A and B.
6. Specify the confidence level Alpha.
7. Specify a location for the output, and click [OK].

Example XL9–4

At $\alpha = 0.05$, test the hypothesis that the two population variances are equal, using the sample data provided here.

Set A	63	73	80	60	86	83	70	72	82
Set B	86	93	64	82	81	75	88	63	63

The results appear in the table that Excel generates, shown here. For this example, the output shows that the null hypothesis cannot be rejected at an α level of 0.05.



F-Test Two-Sample for Variances

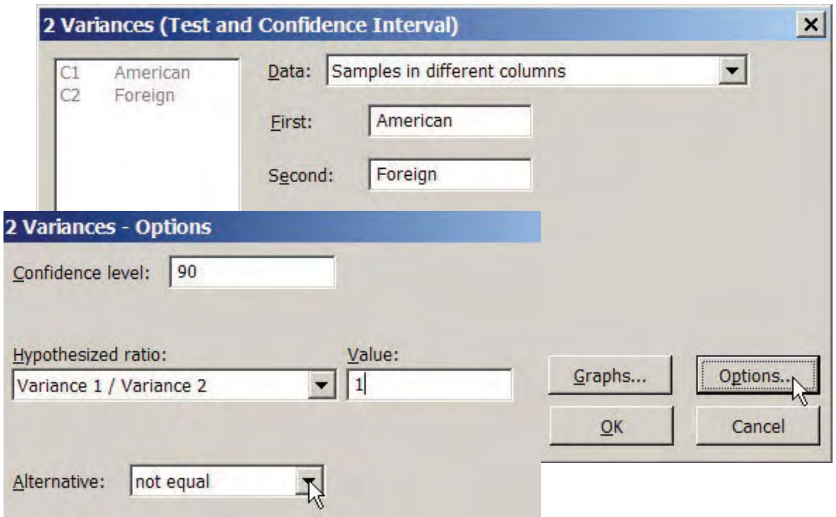
	Variable 1	*Variable 2*
Mean	74.33333333	77.22222222
Variance	82.75	132.9444444
Observations	9	9
df	8	8
F	0.622440451	
P(F<=f) one-tail	0.258814151	
F Critical one-tail	0.290858219	

MINITAB
Step by Step

Test for the Difference Between Two Variances

For Example 9–16, test the hypothesis that the variance in the number of passengers for American and foreign airports is different. Use the *P*-value approach.

American airports	**Foreign airports**
36.8	60.7
72.4	42.7
60.5	51.2
73.5	38.6
61.2	
40.1	



2 Variances (Test and Confidence Interval)

C1 American
C2 Foreign

Data: Samples in different columns

First: American

Second: Foreign

2 Variances - Options

Confidence level: 90

Hypothesized ratio: Variance 1 / Variance 2

Value: 1

Alternative: not equal

1. Enter the data into two columns of MINITAB.
2. Name the columns American and Foreign.
   a) Select **Stat>Basic Statistics>2-Variances.**
   b) Select Each sample is in its own column from the drop down menu.
   c) Click in the text box for Sample 1, then double-click C1 American.
   d) Double-click C2 Foreign, then click on [Options]. The dialog box is shown. For ratio, select (sample 1 variance) / (sample 2 variance) and change the confidence level to **90**. The hypothesized ratio should be 1. For the Alternative hypothesis, select Ratio > hypothesized ratio. Check the box for Use test and confidence intervals based on normal distribution.
3. Click [OK] twice. A graph window will open that includes a small window that says the *P*-value is 0.234. In the session window, the F-test statistic is shown as the Ratio of variances = 2.570. You can view the session window by closing the graph or clicking and dragging it to the right hand part of your screen.

There is not enough evidence in the sample to conclude there is greater variance in the number of passengers in American airports compared to foreign airports.

Summary

Many times researchers are interested in comparing two parameters such as two means, two proportions, or two variances. These measures are obtained from two samples, then compared using a z test, t test, or an F test.

- If two sample means are compared, when the samples are independent and the population standard deviations are known, a z test is used. If the sample sizes are less than 30, the populations should be normally distributed. (9–1)
- If two means are compared when the samples are independent and the sample standard deviations are used, then a t test is used. The two variances are assumed to be unequal. (9–2)
- When the two samples are dependent or related, such as using the same subjects and comparing the means of before-and-after tests, then the t test for dependent samples is used. (9–3)
- Two proportions can be compared by using the z test for proportions. In this case, each of n_1p_1, n_1q_1, n_2p_2, and n_2q_2 must all be 5 or more. (9–4)
- Two variances can be compared by using an F test. The critical values for the F test are obtained from the F distribution. (9–5)
- Confidence intervals for differences between two parameters can also be found.

Important Terms

dependent samples 507

F distribution 529

F test 528

independent samples 499

pooled estimate of the variance 502

Important Formulas

Formula for the z test for comparing two means from independent populations; σ_1 and σ_2 are known:

$$z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

Formula for the confidence interval for difference of two means when σ_1 and σ_2 are known:

$$(\overline{X}_1 - \overline{X}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\overline{X}_1 - \overline{X}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Formula for the t test for comparing two means (independent samples, variances not equal), σ_1 and σ_2 unknown:

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

and d.f. = the smaller of $n_1 - 1$ or $n_2 - 1$.

Formula for the confidence interval for the difference of two means (independent samples, variances unequal), σ_1 and σ_2 unknown:

$$(\overline{X}_1 - \overline{X}_2) - t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < (\overline{X}_1 - \overline{X}_2) + t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

and d.f. = smaller of $n_1 - 1$ and $n_2 - 1$.

Formula for the t test for comparing two means from dependent samples:

$$t = \frac{\overline{D} - \mu_D}{s_D/\sqrt{n}}$$

d.f. $= n - 1$

where $\overline{D}$ is the mean of the differences

$$\overline{D} = \frac{\Sigma D}{n}$$

and s_D is the standard deviation of the differences

$$s_D = \sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{n(n-1)}}$$

Formula for confidence interval for the mean of the difference for dependent samples:

$$\overline{D} - t_{\alpha/2}\frac{s_D}{\sqrt{n}} < \mu_D < \overline{D} + t_{\alpha/2}\frac{s_D}{\sqrt{n}}$$

and d.f. $= n - 1$.

Formula for the z test for comparing two proportions:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\overline{p}\,\overline{q}\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$$

where

$$\overline{p} = \frac{X_1 + X_2}{n_1 + n_2} \qquad \hat{p}_1 = \frac{X_1}{n_1}$$

$$\overline{q} = 1 - \overline{p} \qquad \hat{p}_2 = \frac{X_2}{n_2}$$

Formula for confidence interval for the difference of two proportions:

$$(\hat{p}_1 - \hat{p}_2) - z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} < p_1 - p_2 < (\hat{p}_1 - \hat{p}_2) + z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

Formula for the F test for comparing two variances:

$$F = \frac{s_1^2}{s_2^2} \qquad \text{d.f.N.} = n_1 - 1 \qquad \text{d.f.D.} = n_2 - 1$$

The larger variance is placed in the numerator.

Review Exercises

For each exercise, perform these steps. Assume that all variables are normally or approximately normally distributed.

a. State the hypotheses and identify the claim.
b. Find the critical value(s).
c. Compute the test value.
d. Make the decision.
e. Summarize the results.

Use the traditional method of hypothesis testing unless otherwise specified.

Section 9–1

1. **Driving for Pleasure** Two groups of randomly selected drivers are surveyed to see how many miles per week they drive for pleasure trips. The data are shown. At $\alpha = 0.01$, can it be concluded that single drivers do more driving for pleasure trips on average than married drivers? Assume $\sigma_1 = 16.7$ and $\sigma_2 = 16.1$.

Single drivers					Married drivers				
106	110	115	121	132	97	104	138	102	115
119	97	118	122	135	133	120	119	136	96
110	117	116	138	142	139	108	117	145	114
115	114	103	98	99	140	136	113	113	150
108	117	152	147	117	101	114	116	113	135
154	86	115	116	104	115	109	147	106	88
107	133	138	142	140	113	119	99	108	105

2. **Average Earnings of College Graduates** The average yearly earnings of male college graduates (with at least a bachelor's degree) are $58,500 for men aged 25 to 34. The average yearly earnings of female college graduates with the same qualifications are $49,339. Based on the results below, can it be concluded that there is a difference in mean earnings between male and female college graduates? Use the 0.01 level of significance.

	Male	**Female**
Sample mean	$59,235	$52,487
Population standard deviation	$8,945	$10,125
Sample size	40	35

Source: New York Times Almanac.

Section 9–2

3. **Physical Therapy** A recent study of 20 individuals found that the average number of therapy sessions a person takes for a shoulder problem is 9.6. The standard deviation of the sample was 2.8. A study of 25 individuals with a hip problem found that they had a mean of 10.3 sessions. The standard deviation for this sample was 2.3. At $\alpha = 0.01$, is there a significant difference in the means?

4. **Average Temperatures** The average temperatures for a 25-day period for Birmingham, Alabama, and Chicago, Illinois, are shown. Based on the samples, at $\alpha = 0.10$, can it be concluded that it is warmer in Birmingham?

Birmingham					**Chicago**				
78	82	68	67	68	70	74	73	60	77
75	73	75	64	68	71	72	71	74	76
62	73	77	78	79	71	80	65	70	83
74	72	73	78	68	67	76	75	62	65
73	79	82	71	66	66	65	77	66	64

5. **Teachers' Salaries** A random sample of 15 teachers from Rhode Island has an average salary of $35,270, with a standard deviation of $3256. A random sample of 30 teachers from New York has an average salary of $29,512, with a standard deviation of $1432. Is there a significant difference in teachers' salaries between the two states? Use $\alpha = 0.02$. Find the 98% confidence interval for the difference of the two means.

6. **Soft Drinks in School** The data show the amounts (in thousands of dollars) of the contracts for soft drinks in randomly selected local school districts. At $\alpha = 0.10$, can it be concluded that there is a difference in the averages? Use the *P*-value method. Give a reason why the result would be of concern to a cafeteria manager.

Pepsi						**Coca-Cola**		
46	120	80	500	100	59	420	285	57

Source: Local school districts.

Section 9–3

7. **High and Low Temperatures** March is a month of variable weather in the Northeast. The chart shows records of the actual high and low temperatures for a selection of days in March from the weather report for Pittsburgh, Pennsylvania. At the 0.01 level of significance, is there sufficient evidence to conclude that there is more than a 10° difference between average highs and lows?

Maximum	44	46	46	36	34	36	57	62	73	53
Minimum	27	34	24	19	19	26	33	57	46	26

Source: www.wunderground.com

8. **Testing After Review** A statistics class was given a pretest on probability (since many had previous experience in some other class). Then the class was given a six-page review handout to study for two days. At the next class they were given another test. Is there sufficient evidence that the scores improved? Use $\alpha = 0.05$.

Student	1	2	3	4	5	6
Pretest	52	50	40	58	60	52
Posttest	62	65	50	65	68	63

Section 9–4

9. **Lay Teachers in Religious Schools** A study found a slightly lower percentage of lay teachers in religious secondary schools than in elementary schools. A random sample of 200 elementary school and 200 secondary school teachers from religious schools in a large diocese found the following. At the 0.05 level of significance, is there sufficient evidence to conclude a difference in proportions?

	Elementary	**Secondary**
Sample size	200	200
Lay teachers	49	62

Source: New York Times Almanac.

10. **Gambling** A survey of 60 men found that 36 gamble. Another survey of 50 women found that 28 gamble. At $\alpha = 0.01$, is there a difference in the proportions?

Section 9–5

11. **Noise Levels in Hospitals** In the hospital study cited previously, the standard deviation of the noise levels of the 11 intensive care units was 4.1 dBA, and the standard deviation of the noise levels of 24 nonmedical care areas, such as kitchens and machine rooms, was 13.2 dBA. At $\alpha = 0.10$, is there a significant difference between the standard deviations of these two areas?

Source: M. Bayo, A. Garcia, and A. Garcia, "Noise Levels in an Urban Hospital and Workers' Subjective Responses," *Archives of Environmental Health.*

12. **Heights of World Famous Cathedrals** The heights (in feet) for a random sample of world famous cathedrals are listed. In addition, the heights for a random sample of the tallest buildings in the world are listed. Is there sufficient

evidence at $\alpha = 0.05$ to conclude that there is a difference in the variances in height between the two groups?

Cathedrals	72	114	157	56	83	108	90	151	
Tallest buildings	452	442	415	391	355	344	310	302	209

Source: www.infoplease.com

13. **Sodium Content of Cereals** The sodium content of brands of cereal produced by two major manufacturers is shown. At $\alpha = 0.01$, is there a significant difference in the variances?

Manufacturer 1			**Manufacturer 2**		
87	92	96	87	92	93
100	94	94	91	100	94
101	103	98	103	96	98
91	92	96	87	92	91

STATISTICS TODAY

To Vaccinate or Not to Vaccinate? Small or Large? —Revisited

Using a *z* test to compare two proportions, the researchers found that the proportion of residents in smaller nursing homes who were vaccinated (80.8%) was statistically greater than that of residents in large nursing homes who were vaccinated (68.7%). Using statistical methods presented in later chapters, they also found that the larger size of the nursing home and the lower frequency of vaccination were significant predictions of influenza outbreaks in nursing homes.

Data Analysis

The Data Bank is found in Appendix B, or on the World Wide Web by following links from www.mhhe.com/math/stat/bluman/

1. From the Data Bank, select a variable and compare the mean of the variable for a random sample of at least 30 men with the mean of the variable for the random sample of at least 30 women. Use a *z* test.
2. Repeat the experiment in Exercise 1, using a different variable and two samples of size 15. Compare the means by using a *t* test.
3. Compare the proportion of men who are smokers with the proportion of women who are smokers. Use the data in the Data Bank. Choose random samples of size 30 or more. Use the *z* test for proportions.
4. Select two samples of 20 values from the data in Data Set IV in Appendix B. Test the hypothesis that the mean heights of the buildings are equal.
5. Using the same data obtained in Exercise 4, test the hypothesis that the variances are equal.

Chapter Quiz

Determine whether each statement is true or false. If the statement is false, explain why.

1. When you are testing the difference between two means, it is not important to distinguish whether the samples are independent of each other.
2. If the same diet is given to two groups of randomly selected individuals, the samples are considered to be dependent.
3. When computing the *F* test value, you should place the larger variance in the numerator of the fraction.
4. Tests for variances are always two-tailed.

Select the best answer.

5. To test the equality of two variances, you would use a(n) _______ test.

   a. *z* *c.* Chi-square
   b. *t* *d.* *F*

6. To test the equality of two proportions, you would use a(n) _______ test.

   a. *z* *c.* Chi-square
   b. *t* *d.* *F*

7. The mean value of *F* is approximately equal to

   a. 0 *c.* 1
   b. 0.5 *d.* It cannot be determined.

8. What test can be used to test the difference between two sample means when the population variances are known?

   a. *z* *c.* Chi-square
   b. *t* *d.* *F*

Complete these statements with the best answer.

9. If you hypothesize that there is no difference between means, this is represented as H_0: ________.

10. When you are testing the difference between two means, the ________ test is used when the population variances are not known.

11. When the t test is used for testing the equality of two means, the populations must be ________.

12. The values of F cannot be ________.

13. The formula for the F test for variances is ________.

For each of these problems, perform the following steps.

a. State the hypotheses and identify the claim.
b. Find the critical value(s).
c. Compute the test value.
d. Make the decision.
e. Summarize the results.

Use the traditional method of hypothesis testing unless otherwise specified.

14. Cholesterol Levels A researcher wishes to see if there is a difference in the cholesterol levels of two groups of men. A random sample of 30 men between the ages of 25 and 40 is selected and tested. The average level is 223. A second random sample of 25 men between the ages of 41 and 56 is selected and tested. The average of this group is 229. The population standard deviation for both groups is 6. At $\alpha = 0.01$, is there a difference in the cholesterol levels between the two groups? Find the 99% confidence interval for the difference of the two means.

15. Apartment Rental Fees The data shown are the rental fees (in dollars) for two random samples of apartments in a large city. At $\alpha = 0.10$, can it be concluded that the average rental fee for apartments in the east is greater than the average rental fee in the west? Assume $\sigma_1 = 119$ and $\sigma_2 = 103$.

East					West				
495	390	540	445	420	525	400	310	375	750
410	550	499	500	550	390	795	554	450	370
389	350	450	530	350	385	395	425	500	550
375	690	325	350	799	380	400	450	365	425
475	295	350	485	625	375	360	425	400	475
275	450	440	425	675	400	475	430	410	450
625	390	485	550	650	425	450	620	500	400
685	385	450	550	425	295	350	300	360	400

Source: Pittsburgh Post-Gazette.

16. Prices of Low-Calorie Foods The average price of a random sample of 12 bottles of diet salad dressing taken from different stores is \$1.43. The standard deviation is \$0.09. The average price of a random sample of 16 low-calorie frozen desserts is \$1.03. The standard deviation is \$0.10. At $\alpha = 0.01$, is there a significant difference in price? Find the 99% confidence interval of the difference in the means.

17. Jet Ski Accidents The data shown represent the number of accidents people had when using jet skis and other types of wet bikes. At $\alpha = 0.05$, can it be concluded that the average number of accidents per year has increased from one period to the next?

Earlier period			Later period		
376	650	844	1650	2236	3002
1162	1513		4028	4010	

Source: USA TODAY.

18. Salaries of Chemists A random sample of 12 chemists from Washington state shows an average salary of \$39,420 with a standard deviation of \$1659, while a random sample of 26 chemists from New Mexico has an average salary of \$30,215 with a standard deviation of \$4116. Is there a significant difference between the two states in chemists' salaries at $\alpha = 0.02$? Find the 98% confidence interval of the difference in the means.

19. Family Incomes The average income of 15 randomly selected families who reside in a large metropolitan East Coast city is \$62,456. The standard deviation is \$9652. The average income of 11 randomly selected families who reside in a rural area of the Midwest is \$60,213, with a standard deviation of \$2009. At $\alpha = 0.05$, can it be concluded that the families who live in the cities have a higher income than those who live in the rural areas? Use the P-value method.

20. Mathematical Skills In an effort to improve the mathematical skills of 10 students, a teacher provides a weekly 1-hour tutoring session for the students. A pretest is given before the sessions, and a posttest is given after. The results are shown here. At $\alpha = 0.01$, can it be concluded that the sessions help to improve the students' mathematical skills?

Student	1	2	3	4	5	6	7	8	9	10
Pretest	82	76	91	62	81	67	71	69	80	85
Posttest	88	80	98	80	80	73	74	78	85	93

21. Egg Production To increase egg production, a farmer decided to increase the amount of time the lights in his hen house were on. Ten hens were randomly selected, and the number of eggs each produced was recorded. After one week of lengthened light time, the same hens were monitored again. The data are given here. At $\alpha = 0.05$, can it be concluded that the increased light time increased egg production?

Hen	1	2	3	4	5	6	7	8	9	10
Before	4	3	8	7	6	4	9	7	6	5
After	6	5	9	7	4	5	10	6	9	6

22. Factory Worker Literacy Rates In a random sample of 80 workers from a factory in city A, it was found that 5% were unable to read, while in a random sample of 50 workers in city B, 8% were unable to read. Can it be concluded that there is a difference in the proportions of nonreaders in the two cities? Use $\alpha = 0.10$. Find the 90% confidence interval for the difference of the two proportions.

23. Male Head of Household A recent survey of 200 randomly selected households showed that 8 had a single male as the head of household. Forty years ago, a survey of 200 randomly selected households showed that 6 had a single male as the head of household. At $\alpha = 0.05$, can it be concluded that the proportion has changed? Find the 95% confidence interval of the difference of the two proportions. Does the confidence interval contain 0? Why is this important to know?

Source: Based on data from the U.S. Census Bureau.

24. Money Spent on Road Repair A politician wishes to compare the variances of the amount of money spent for road repair in two different counties. The data are given here. At $\alpha = 0.05$, is there a significant difference in the variances of the amounts spent in the two counties? Use the *P*-value method.

County A	County B
$s_1 = \$11{,}596$	$s_2 = \$14{,}837$
$n_1 = 15$	$n_2 = 18$

25. Heights of Basketball Players A researcher wants to compare the variances of the heights (in inches) of four-year college basketball players with those of players in junior colleges. A random sample of 30 players from each type of school is selected, and the variances of the heights for each type are 2.43 and 3.15, respectively. At $\alpha = 0.10$, is there a significant difference between the variances of the heights in the two types of schools?

Data Projects

Use a significance level of 0.05 for all tests below.

1. **Business and Finance** Use the data collected in data project 1 of Chapter 2 to complete this problem. Test the claim that the mean earnings per share for Dow Jones stocks are greater than for NASDAQ stocks.

2. **Sports and Leisure** Use the data collected in data project 2 of Chapter 7 regarding home runs for this problem. Test the claim that the mean number of home runs hit by the American League sluggers is the same as the mean for the National League.

3. **Technology** Use the cell phone data collected for data project 2 in Chapter 8 to complete this problem. Test the claim that the mean length for outgoing calls is the same as that for incoming calls. Test the claim that the standard deviation for outgoing calls is more than that for incoming calls.

4. **Health and Wellness** Use the data regarding BMI that were collected in data project 6 of Chapter 7 to complete this problem. Test the claim that the mean BMI for males is the same as that for females. Test the claim that the standard deviation for males is the same as that for females.

5. **Politics and Economics** Using data from the Internet for the last Presidential election to categorize the 50 states as “red” or “blue” based on who was supported for President in that state, the Democratic or Republican candidate, test the claim that the mean incomes for red states and blue states are equal.

6. **Your Class** Use the data collected in data project 6 of Chapter 2 regarding heart rates. Test the claim that the heart rates after exercise are more variable than the heart rates before exercise.

Answers to Applying the Concepts

Section 9–1 Home Runs

1. The population of all home runs hit by major league baseball players.

2. A cluster sample was used.

3. Answers will vary. While this sample is not representative of all major league baseball players per se, it does allow us to compare the leaders in each league.

4. H_0: $\mu_1 = \mu_2$ and H_1: $\mu_1 \neq \mu_2$

5. Answers will vary. Possible answers include the 0.05 and 0.01 significance levels.

6. We will use the z test for the difference in means.

7. Our test statistic is $z = \dfrac{44.75 - 42.88}{\sqrt{\dfrac{8.8^2}{40} + \dfrac{7.8^2}{40}}} = 1.01$, and our *P*-value is 0.3124.

8. We fail to reject the null hypothesis.

9. There is not enough evidence to conclude that there is a difference in the number of home runs hit by National League versus American League baseball players.

10. Answers will vary. One possible answer is that since we do not have a random sample of data from each league, we cannot answer the original question asked.

11. Answers will vary. One possible answer is that we could get a random sample of data from each league from a recent season.

Section 9–2 Too Long on the Telephone

1. These samples are independent.

2. We compare the *P*-value of 0.06317 to the significance level to check if the null hypothesis should be rejected.

3. The P-value of 0.06317 also gives the probability of a type I error.
4. Since two critical values are shown, we know that a two-tailed test was done.
5. Since the P-value of 0.06317 is greater than the significance value of 0.05, we fail to reject the null hypothesis and find that we do not have enough evidence to conclude that there is a difference in the lengths of telephone calls made by employees in the two divisions of the company.
6. If the significance level had been 0.10, we would have rejected the null hypothesis, since the P-value would have been less than the significance level.

Section 9–3 Air Quality

1. The purpose of the study is to determine if the air quality in the United States has changed over the past 2 years.
2. These are dependent samples, since we have two readings from each of 10 metropolitan areas.
3. The hypotheses we will test are H_0: $\mu_D = 0$ and H_1: $\mu_D \neq 0$.
4. We will use the 0.05 significance level and critical values of $t = \pm 2.262$.
5. We will use the t test for dependent samples.
6. There are $10 - 1 = 9$ degrees of freedom.
7. Our test statistic is $t = \dfrac{-6.7 - 0}{11.27/\sqrt{10}} = -1.879$. We fail to reject the null hypothesis and find that there is not enough evidence to conclude that the air quality in the United States has changed over the past 2 years.
8. No, we could not use an independent means test since we have two readings from each metropolitan area.
9. Answers will vary. One possible answer is that there are other measures of air quality that we could have examined to answer the question.

Section 9–4 Smoking and Education

1. Our hypotheses are H_0: $p_1 = p_2$ and H_1: $p_1 \neq p_2$.
2. At the 0.05 significance level, our critical values are $z = \pm 1.96$.
3. We will use the z test for the difference between proportions.
4. To complete the statistical test, we would need the sample sizes.
5. Knowing the sample sizes were 1000, we can now complete the test.
6. Our test statistic is

$$z = \frac{0.323 - 0.145}{\sqrt{(0.234)(0.766)\left(\dfrac{1}{1000} + \dfrac{1}{1000}\right)}} = 9.40,$$

   and our P-value is very close to zero. We reject the null hypothesis and find that there is enough evidence to conclude that there is a difference in the proportions of public school students and private school students who smoke.

Section 9–5 Variability and Automatic Transmissions

1. The null hypothesis is that the variances are the same: H_0: $\sigma_1^2 = \sigma_2^2$ (H_1: $\sigma_1^2 \neq \sigma_2^2$).
2. We will use an F test.
3. The value of the test statistic is $F = \dfrac{s_1^2}{s_2^2} = \dfrac{18{,}163.58^2}{17{,}400.57^2} =$ 1.090 and the P-value > 0.05. There is not a significant difference in the variability of the prices between the two countries.
4. Small sample sizes are highly impacted by outliers.
5. The degrees of freedom for the numerator and denominator are both 3.
6. Yes, two sets of data can center on the same mean but have very different standard deviations.

Hypothesis-Testing Summary 1

1. Comparison of a sample mean with a specific population mean.

   Example: H_0: $\mu = 100$

   a. Use the z test when σ is known:

$$z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

   b. Use the t test when σ is unknown:

$$t = \frac{\overline{X} - \mu}{s/\sqrt{n}} \quad \text{with d.f.} = n - 1$$

2. Comparison of a sample variance or standard deviation with a specific population variance or standard deviation.

   Example: H_0: $\sigma^2 = 225$

   Use the chi-square test:

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \quad \text{with d.f.} = n - 1$$

3. Comparison of two sample means.

   Example: H_0: $\mu_1 = \mu_2$

   a. Use the z test when the population variances are known:

$$z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

b. Use the *t* test for independent samples when the population variances are unknown and assume the sample variances are unequal:

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

with d.f. = the smaller of $n_1 - 1$ or $n_2 - 1$.

c. Use the *t* test for means for dependent samples:

Example: H_0: $\mu_D = 0$

$$t = \frac{\overline{D} - \mu_D}{s_D/\sqrt{n}} \qquad \text{with d.f.} = n - 1$$

where n = number of pairs.

4. Comparison of a sample proportion with a specific population proportion.

Example: H_0: $p = 0.32$

Use the *z* test:

$$z = \frac{X - \mu}{\sigma} \qquad \text{or} \qquad z = \frac{\hat{p} - p}{\sqrt{pq/n}}$$

5. Comparison of two sample proportions.

Example: H_0: $p_1 = p_2$

Use the *z* test:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\overline{p}\,\overline{q}\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$$

where

$$\overline{p} = \frac{X_1 + X_2}{n_1 + n_2} \qquad \hat{p}_1 = \frac{X_1}{n_1}$$

$$\overline{q} = 1 - \overline{p} \qquad \hat{p}_2 = \frac{X_2}{n_2}$$

6. Comparison of two sample variances or standard deviations.

Example: H_0: $\sigma_1^2 = \sigma_2^2$

Use the *F* test:

$$F = \frac{s_1^2}{s_2^2}$$

where

s_1^2 = larger variance d.f.N. = $n_1 - 1$

s_2^2 = smaller variance d.f.D. = $n_2 - 1$

11

Other Chi-Square Tests

© Mitch Hrdlicka/Getty Images RF

OUTLINE

Introduction
11–1 Test for Goodness of Fit
11–2 Tests Using Contingency Tables
Summary

OBJECTIVES

After completing this chapter, you should be able to:

1. Test a distribution for goodness of fit, using chi-square.
2. Test two variables for independence, using chi-square.
3. Test proportions for homogeneity, using chi-square.

STATISTICS TODAY

Statistics and Heredity

An Austrian monk, Gregor Mendel (1822–1884), studied genetics, and his principles are the foundation for modern genetics. Mendel used his spare time to grow a variety of peas at the monastery. One of his many experiments involved crossbreeding peas that had smooth yellow seeds with peas that had wrinkled green seeds. He noticed that the results occurred with regularity. That is, some of the offspring had smooth yellow seeds, some had smooth green seeds, some had wrinkled yellow seeds, and some had wrinkled green seeds. Furthermore, after several experiments, the percentages of each type seemed to remain approximately the same. Mendel formulated his theory based on the assumption of dominant and recessive traits and tried to predict the results. He then crossbred his peas and examined 556 seeds over the next generation.

Finally, he compared the actual results with the theoretical results to see if his theory was correct. To do this, he used a "simple" chi-square test, which is explained in this chapter. See Statistics Today—Revisited at the end of this chapter.

Introduction

The chi-square distribution was used in Chapters 7 and 8 to find a confidence interval for a variance or standard deviation and to test a hypothesis about a single variance or standard deviation.

It can also be used for tests concerning *frequency distributions,* such as "If a sample of buyers is given a choice of automobile colors, will each color be selected with the same frequency?" The chi-square distribution can be used to test the *independence* of two variables, for example, "Are senators' opinions on gun control independent of party affiliations?" That is, do the Republicans feel one way and the Democrats feel differently, or do they have the same opinion?

Finally, the chi-square distribution can be used to test the *homogeneity of proportions.* For example, is the proportion of high school seniors who attend college immediately after graduating the same for the northern, southern, eastern, and western parts of the United States?

This chapter explains the chi-square distribution and its applications. In addition to the applications mentioned here, chi-square has many other uses in statistics.

11–1 Test for Goodness of Fit

OBJECTIVE ❶

Test a distribution for goodness of fit, using chi-square.

In addition to being used to test a single variance, the chi-square statistic can be used to see whether a frequency distribution fits a specific pattern. For example, to meet customer demands, a manufacturer of running shoes may wish to see whether buyers show a preference for a specific style. A traffic engineer may wish to see whether accidents occur more often on some days than on others, so that she can increase police patrols accordingly. An emergency service may want to see whether it receives more calls at certain times of the day than at others, so that it can provide adequate staffing.

Recall the characteristics of the chi-square distribution:

1. The chi-square distribution is a family of curves based on the degrees of freedom.
2. The chi-square distributions are positively skewed.
3. All chi-square values are greater than or equal to zero.
4. The total area under each chi-square distribution is equal to 1.

Historical Note

Karl Pearson (1857–1936) first used the chi-square distribution as a goodness-of-fit test for data. He developed many types of descriptive graphs and gave them unusual names such as stigmograms, topograms, stereograms, and radiograms.

When you are testing to see whether a frequency distribution fits a specific pattern, you can use the chi-square *goodness-of-fit test.*

The **chi-square goodness-of-fit test** is used to test the claim that an observed frequency distribution fits some given expected frequency distribution.

For example, suppose you wanted to see if there was a difference in the number of arrests in a certain city for four types of crimes. A random sample of 160 arrests showed the following distribution.

Larceny thefts	Property crimes	Drug use	Driving under the influence
38	50	28	44

Since the frequencies for each flavor were obtained from a sample, these actual frequencies are called the **observed frequencies.** The frequencies obtained by calculation (as if there were no preference) are called the **expected frequencies.**

To calculate the expected frequencies, there are two rules to follow.

1. If all the expected frequencies are equal, the expected frequency E can be calculated by using $E = n/k$, where n is the total number of observations and k is the number of categories.

2. If all the expected frequencies are not equal, then the expected frequency E can be calculated by $E = n \cdot p$, where n is the total number of observations and p is the probability for that category.

Looking at the number of arrests example, if there were no difference, you would expect $160 \div 4 = 40$ arrests for each category. That is, *approximately* 40 people would be arrested for each type of crime. A completed table is shown.

	Larceny thefts	Property crimes	Drug use	Driving under the influence
Observed	38	50	28	44
Expected	40	40	40	40

The observed frequencies will almost always differ from the expected frequencies due to sampling error; that is, the values differ from sample to sample. But the question is: Are these differences significant (there is a difference in the number of arrests for these types of crimes) or are they due to chance? The chi-square goodness-of-fit test will enable the researcher to determine the answer.

Before computing the test value, you must state the hypotheses. The null hypothesis should be a statement indicating that there is no difference or no change. For this example, the hypotheses are as follows:

H_0: There is no difference in the number of arrests for each type of crime.

H_1: There is a difference in the number of arrests for each type of crime.

Next, we need a measure of discrepancy between the observed values O and the expected values E, so we use the test statistic for the chi-square goodness-of-fit test.

Interesting Fact

Men begin to lose their hearing more than 30 years before women. The difference may be due to males' more frequent exposure to such noisy machines as power tools and lawnmowers.

Formula for the Chi-Square Goodness-of-Fit Test

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

with degrees of freedom equal to the number of categories minus 1, and where

O = observed frequency

E = expected frequency

Notice that the value of the test statistic is based on the difference between the observed values and the expected values. If the observed values are significantly different from the expected values, then there is enough evidence to reject the null hypothesis.

When there is perfect agreement between the observed and the expected values, $\chi^2 = 0$. Also, χ^2 can never be negative. Finally, the test is right-tailed because "H_0: Good fit" and "H_1: Not a good fit" mean that χ^2 will be small in the first case and large in the second case.

In the goodness-of-fit test, the degrees of freedom are equal to the number of categories minus 1. In this example, there are four categories; hence, the degrees of freedom are $4 - 1 = 3$. This is so because the number of subjects in each of the first three categories is free to vary. But in order for the sum to be 160, the total number of subjects in the last category is fixed.

Two assumptions are needed for the goodness-of-fit test. These assumptions are given next.

Assumptions for the Chi-Square Goodness-of-Fit Test

1. The data are obtained from a random sample.
2. The expected frequency for each category must be 5 or more.

In this book, the assumptions will be stated in the exercises; however, when encountering statistics in other situations, you must check to see that these assumptions have been met before proceeding.

For use of the chi-square goodness-of-fit test, statisticians have determined that the expected frequencies should be at least 5, as stated in the assumptions. The reasoning is as follows: The chi-square distribution is continuous, whereas the goodness-of-fit test is discrete. However, the continuous distribution is a good approximation and can be used when the expected value for each class is at least 5. If an expected frequency of a class is less than 5, then that class can be combined with another class so that the expected frequency is 5 or more.

The steps for the chi-square goodness-of-fit test are summarized in this Procedure Table.

Procedure Table

The Chi-Square Goodness-of-Fit Test

Step 1 State the hypotheses and identify the claim.

Step 2 Find the critical value from Table G. The test is always right-tailed.

Step 3 Compute the test value.

Find the sum of the $\frac{(O-E)^2}{E}$ values.

Step 4 Make the decision.

Step 5 Summarize the results.

Example 11–1 demonstrates the situation when the expected values of the categories are equal.

EXAMPLE 11–1 Arrests for Crimes

Is there enough evidence to reject the claim that the number of arrests for each category of crimes is the same? Use $\alpha = 0.05$.

SOLUTION

Step 1 State the hypotheses and identify the claim.

H_0: There is no difference in the number of arrests for each type of crime. (claim)
H_1: There is a difference in the number of arrests for each type of crime.

Step 2 Find the critical value. The degrees of freedom are $4 - 1 = 3$, and at $\alpha = 0.05$ the critical value from Table G in Appendix A is 7.815.

Step 3 Compute the test value. Note the expected values are found by $E = n/k = 160/4 = 40$.

The table looks like this.

	Larceny thefts	Property crimes	Drug use	Driving under the influence
Observed	38	50	28	44
Expected	40	40	40	40

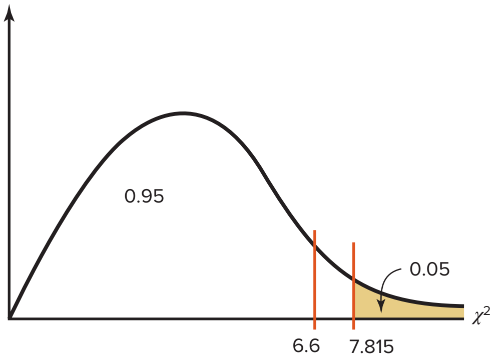
The test value is computed by subtracting the expected value corresponding to the observed value, squaring the result, and dividing by the expected value. Then find the sum of these values.

$$\chi^2 = \sum\frac{(0-E)^2}{E} = \frac{(38-40)^2}{40} + \frac{(50-40)^2}{40} + \frac{(28-40)^2}{40} + \frac{(44-40)^2}{40}$$
$$= 0.1 + 2.5 + 3.6 + 0.4$$
$$= 6.6$$

Step 4 Make the decision. The decision is to not reject the null hypothesis since $6.6 < 7.815$, as shown in Figure 11–1.

FIGURE 11–1

Critical and Test Values for Example 11–1



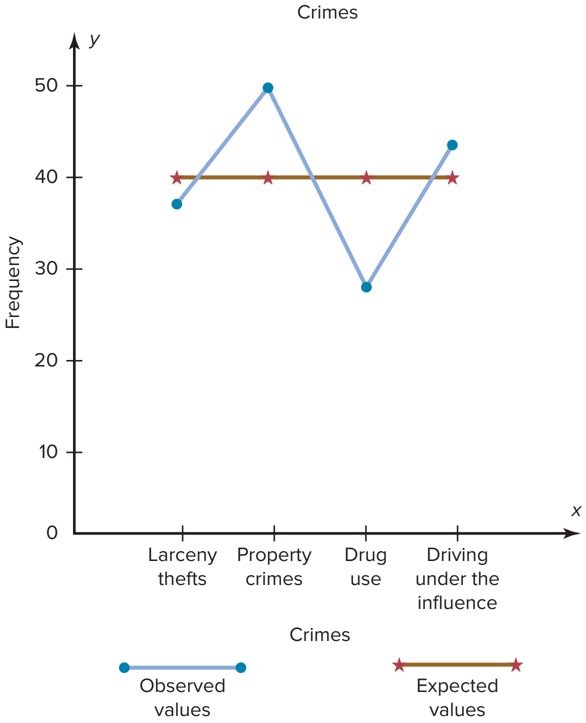
Step 5 Summarize the results. There is not enough evidence to reject the claim that there is no difference in the number of arrests for each type of crime.

Also the P-values can be found for this test. In Example 11–1, the test value was 6.6. If you look across the row with d.f. = 3 of Table G, you will find that 6.6 is between 6.251 and 7.815, that is, between 0.10 and 0.05 at the top of the table. This corresponds to $0.05 < p < 0.10$. Since the P-value is greater than 0.05, the decision is to not reject the null hypothesis.

To get some idea of why this test is called the goodness-of-fit test, examine graphs of the observed values and expected values from Example 11–1. See Figure 11–2. From the graphs, you can see whether the observed values and expected values are close together or far apart. In this case, the observed values and the expected values are far apart, so this is not a good fit. As a result, we reject the null hypothesis.

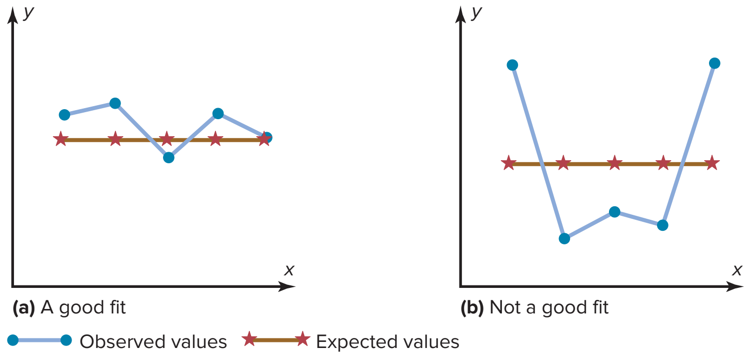
FIGURE 11–2

Graphs of the Observed and Expected Values for Arrests



When the observed values and expected values are close together, the chi-square test value will be small. Then the decision will be to not reject the null hypothesis—hence, this is "a good fit." See Figure 11–3(a). When the observed values and the expected values are far apart, the chi-square test value will be large. Then the null hypothesis will be rejected—hence, there is "not a good fit." See Figure 11–3(b).

FIGURE 11–3
Results of the Goodness-of-Fit Test



(a) A good fit **(b)** Not a good fit

EXAMPLE 11–2 Education Level of Adults

The Census Bureau of the U.S. government found that 13% of adults did not finish high school, 30% graduated from high school only, 29% had some college education but did not obtain a bachelor's degree, and 28% were college graduates. To see if these proportions were consistent with those people who lived in the Lincoln County area, a local researcher selected a random sample of 300 adults and found that 43 did not finish high school, 76 were high school graduates only, 96 had some college education, and 85 were college graduates. At $\alpha = 0.10$, test the claim that the proportions are the same for the adults in Lincoln County as those stated by the Census Bureau.

SOLUTION

Step 1 State the hypotheses and identify the claim.

H_0: The proportion of people in each category is as follows: 13% did not finish high school, 30% were high school graduates only, 29% had some college education but did not graduate, and 28% had a college degree (claim).

H_1: The distribution is not the same as stated in the null hypothesis.

Step 2 Find the critical value. Since $\alpha = 0.10$ and the degrees of freedom are $4 - 1 = 3$, the critical value is 6.251.

Step 3 Compute the test value. First, we must calculate the expected values. Multiply the total number of people surveyed (300) by the percentages of people in each category.

$$0.13 \times 300 = 39$$
$$0.30 \times 300 = 90$$
$$0.29 \times 300 = 87$$
$$0.28 \times 300 = 84$$

The table looks like this:

Frequency	Did not finish high school	H.S. graduate	Some college	College graduate
Observed	43	76	96	85
Expected	39	90	87	84

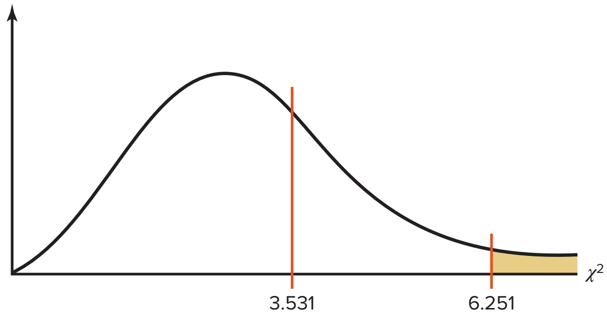
Next calculate the test value.

$$x^2 = \sum \frac{(O - E)^2}{E} = \frac{(43 - 39)^2}{39} + \frac{(76 - 90)^2}{90} + \frac{(96 - 87)^2}{87} + \frac{(85 - 84)^2}{84}$$
$$= 0.410 + 2.178 + 0.931 + 0.012 = 3.531$$

Step 4 Make the decision. Since $3.531 < 6.251$, the decision is not to reject the null hypothesis. See Figure 11–4.

FIGURE 11–4

Critical and Test Values for Example 11–2



Step 5 Summarize the results. There is not enough evidence to reject the claim. It can be concluded that the percentages are not significantly different from those given in the null hypothesis. That is, the proportions are not significantly different from those stated by the U.S. Census Bureau.

EXAMPLE 11–3 Firearm Deaths

A researcher read that firearm-related deaths for people aged 1 to 18 years were distributed as follows: 74% were accidental, 16% were homicides, and 10% were suicides. In her district, there were 68 accidental deaths, 27 homicides, and 5 suicides during the past year. At $\alpha = 0.10$, test the claim that the percentages are equal.

Source: Centers for Disease Control and Prevention.

SOLUTION

Step 1 State the hypotheses and identify the claim:

H_0: The deaths due to firearms for people aged 1 through 18 years are distributed as follows: 74% accidental, 16% homicides, and 10% suicides (claim).

H_1: The distribution is not the same as stated in the null hypothesis.

Step 2 Find the critical value. Since $\alpha = 0.10$ and the degrees of freedom are $3 - 1 = 2$, the critical value is 4.605.

Step 3 Compute the test value.

First calculate the expected values, using the formula $E = n \cdot p$ as shown.

$$100 \times 0.74 = 74$$
$$100 \times 0.16 = 16$$
$$100 \times 0.10 = 10$$

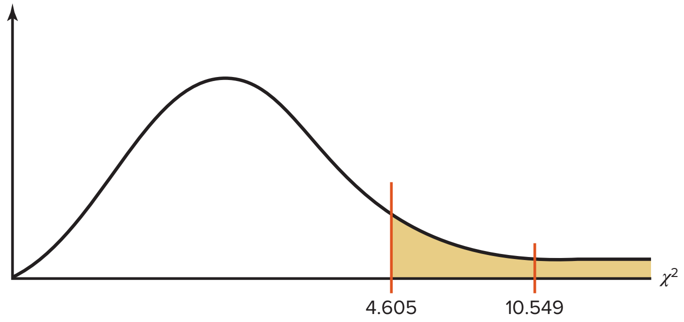
The table looks like this.

Frequency	Accidental	Homicides	Suicides
Observed	68	27	5
Expected	74	16	10

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$
$$= \frac{(68 - 74)^2}{74} + \frac{(27 - 16)^2}{16} + \frac{(5 - 10)^2}{10}$$
$$= 10.549$$

Step 4 Reject the null hypothesis, since 10.549 > 4.605, as shown in Figure 11–5.

FIGURE 11–5
Critical and Test Values for Example 11–3



Step 5 Summarize the results. There is enough evidence to reject the claim that the distribution is 74% accidental, 16% homicides, and 10% suicides.

Test of Normality (Optional)

The chi-square goodness-of-fit test can be used to test a variable to see if it is normally distributed. The null hypotheses are

H_0: The variable is normally distributed.
H_1: The variable is not normally distributed.

The procedure is somewhat complicated. It involves finding the expected frequencies for each class of a frequency distribution by using the standard normal distribution. Then the actual frequencies (i.e., observed frequencies) are compared to the expected frequencies, using the chi-square goodness-of-fit test. If the observed frequencies are close in value to the expected frequencies, then the chi-square test value will be small and the null hypothesis cannot be rejected. In this case, it can be concluded that the variable is approximately normally distributed.

On the other hand, if there is a large difference between the observed frequencies and the expected frequencies, then the chi-square test value will be larger and the null hypothesis can be rejected. In this case, it can be concluded that the variable is not normally distributed. Example 11–4 illustrates the procedure for the chi-square test of normality. To find the areas in the examples, you might want to review Section 6–2.

Example 11–4 shows how to do the calculations.

EXAMPLE 11–4 Test of Normality

Use chi-square to determine if the variable shown in the frequency distribution is normally distributed. Use $\alpha = 0.05$.

Boundaries	Frequency
89.5–104.5	24
104.5–119.5	62
119.5–134.5	72
134.5–149.5	26
149.5–164.5	12
164.5–179.5	4
	Total = 200

SOLUTION

H_0: The variable is normally distributed.
H_1: The variable is not normally distributed.

First find the mean and standard deviation of the variable. (*Note*: s is used to approximate σ.)

Unusual Stat

Drinking milk may lower your risk of stroke. A 22-year study of men over 55 found that only 4% of men who drank 16 ounces of milk every day suffered a stroke, compared with 8% of the nonmilk drinkers.

Boundaries	f	X_m	$f \cdot X_m$	$f \cdot X_m^2$
89.5–104.5	24	97	2,328	225,816
104.5–119.5	62	112	6,944	777,728
119.5–134.5	72	127	9,144	1,161,288
134.5–149.5	26	142	3,692	524,264
149.5–164.5	12	157	1,884	295,788
164.5–179.5	4	172	688	118,336
	200		24,680	3,103,220

$$\overline{X} = \frac{24{,}680}{200} = 123.4$$

$$s = \sqrt{\frac{200(3{,}103{,}220) - 24{,}680^2}{200(199)}} = \sqrt{290} = 17.03$$

Next find the area under the standard normal distribution, using z values and Table E for each class.

The z score for $x = 104.5$ is found as

$$z = \frac{104.5 - 123.4}{17.03} = -1.11$$

The area for $z < -1.11$ is 0.1335.

The z score for 119.5 is found as

$$z = \frac{119.5 - 123.4}{17.03} = -0.23$$

The area for $-1.11 < z < -0.23$ is $0.4090 - 0.1335 = 0.2755$.

The z score for 134.5 is found as

$$z = \frac{134.5 - 123.4}{17.03} = 0.65$$

The area for $-0.23 < z < 0.65$ is $0.7422 - 0.4090 = 0.3332$.

The z score for 149.5 is found as

$$z = \frac{149.5 - 123.4}{17.03} = 1.53$$

The area for $0.65 < z < 1.53$ is $0.9370 - 0.7422 = 0.1948$.

The z score for 164.5 is found as

$$z = \frac{164.5 - 123.4}{17.03} = 2.41$$

The area for $1.53 < z < 2.41$ is $0.9920 - 0.9370 = 0.0550$.
The area for $z > 2.41$ is $1.0000 - 0.9920 = 0.0080$.

Find the expected frequencies for each class by multiplying the area by 200. The expected frequencies are found by

$$\begin{aligned}0.1335 \cdot 200 &= 26.7\\ 0.2755 \cdot 200 &= 55.1\\ 0.3332 \cdot 200 &= 66.64\\ 0.1948 \cdot 200 &= 38.96\\ 0.0550 \cdot 200 &= 11.0\\ 0.0080 \cdot 200 &= 1.6\end{aligned}$$

Note: Since the expected frequency for the last category is less than 5, it can be combined with the previous category.

The table looks like this.

O	24	62	72	26	16
E	26.7	55.1	66.64	38.96	12.6

Finally, find the chi-square test value using the formula $\chi^2 = \sum \frac{(O-E)^2}{E}$.

$$\chi^2 = \frac{(24-26.7)^2}{26.7} + \frac{(62-55.1)^2}{55.1} + \frac{(72-66.64)^2}{66.64} + \frac{(26-38.96)^2}{38.96} + \frac{(16-12.6)^2}{12.6}$$
$$= 6.797$$

The critical value in this test has the degrees of freedom equal to the number of categories minus 3 since 1 degree of freedom is lost for each parameter that is estimated. In this case, the mean and standard deviation have been estimated, so 2 additional degrees of freedom are needed.

The C.V. with d.f. $= 5 - 3 = 2$ and $\alpha = 0.05$ is 5.991, so the null hypothesis is rejected. Hence, the distribution can be considered not normally distributed.

Note: At $\alpha = 0.01$, the C.V. $= 9.210$ and the null hypothesis would not be rejected. Hence, we could consider that the variable is normally distributed at $\alpha = 0.01$. So it is important to decide which level of significance you want to use prior to conducting the test.

Applying the Concepts 11–1

Skittles Color Distribution

M&M/Mars, the makers of Skittles candies, states that the flavor blend is 20% for each flavor. Skittles is a combination of lemon, lime, orange, strawberry, and grape flavored candies. The following data list the results of four randomly selected bags of Skittles and their flavor blends. Use the data to answer the questions.

	Flavor				
Bag	**Yellow**	**Green**	**Orange**	**Red**	**Purple**
1	14	7	20	10	7
2	17	20	5	5	13
3	4	4	16	13	21
4	17	12	9	16	3
Total	52	43	50	44	44

1. Are the variables quantitative or qualitative?
2. What type of test can be used to compare the observed values to the expected values?
3. Compute a chi-square test on the total values.
4. What hypotheses should you use?
5. What are the degrees of freedom for the test? What is the critical value?
6. What is your conclusion?

See page 643 for the answers.

Exercises 11–1

1. How does the goodness-of-fit test differ from the chi-square variance test?

2. How are the degrees of freedom computed for the goodness-of-fit test?

3. How are the expected values computed for the goodness-of-fit test?

4. When the expected frequency is less than 5 for a specific class, what should be done so that you can use the goodness-of-fit test?

For Exercises 5 through 18, perform these steps.

a. State the hypotheses and identify the claim.
b. Find the critical value.
c. Compute the test value.
d. Make the decision.
e. Summarize the results.

Use the traditional method of hypothesis testing unless otherwise specified. Assume all assumptions are met.

5. Statistics Class Times A professor wishes to see if students show a time preference for statistics classes. A sample of four statistics classes shows the enrollment. At $\alpha = 0.01$, do the students show a time preference for the classes?

Time	8:00 AM	10:00 AM	12:00 PM	2:00 PM
Students	24	35	31	26

6. Blood Types A medical researcher wishes to see if hospital patients in a large hospital have the same blood type distribution as those in the general population. The distribution for the general population is as follows: type A, 20%; type B, 28%; type O, 36%; and type AB = 16%. He selects a random sample of 50 patients and finds the following: 12 have type A blood, 8 have type B, 24 have type O, and 6 have type AB blood.

At $\alpha = 0.10$, can it be concluded that the distribution is the same as that of the general population?

7. Extending the School Year A researcher surveyed 100 randomly selected teachers in a large school district and found that 46 wanted to extend the school year, 42 did not, and 12 had no opinion. At the 0.05 level of significance, is the distribution different from the national distribution where 45% wished to extend the school year, 47% did not want the school year extended, and 8% had no opinion?

8. On-Time Performance by Airlines According to the Bureau of Transportation Statistics, on-time performance by the airlines is described as follows:

Action	**% of Time**
On time	70.8
National Aviation System delay	8.2
Aircraft arriving late	9.0
Other (because of weather and other conditions)	12.0

Records of 200 randomly selected flights for a major airline company showed that 125 planes were on time; 40 were delayed because of weather, 10 because of a National Aviation System delay, and the rest because of arriving late. At $\alpha = 0.05$, do these results differ from the government's statistics?

Source: www.transtats.bts.gov

9. Genetically Modified Food An ABC News poll asked adults whether they felt genetically modified food was safe to eat. Thirty-five percent felt it was safe, 52% felt it was not safe, and 13% had no opinion. A random sample of 120 adults was asked the same question at a local county fair. Forty people felt that genetically modified food was safe, 60 felt that it was not safe, and 20 had no opinion. At the 0.01 level of significance, is there sufficient evidence to conclude that the proportions differ from those reported in the poll?

Source: ABCNews.com Poll, www.pollingreport.com

10. Truck Colors In a recent year, the most popular colors for light trucks were white, 31%; black, 19%; silver 11%;

red 11%; gray 10%; blue 8%; and other 10%. A survey of randomly selected light truck owners in a particular area revealed the following. At $\alpha = 0.05$, do the proportions differ from those stated?

White	Black	Silver	Red	Gray	Blue	Other
45	32	30	30	22	15	6

Source: World Almanac.

11. Employee Absences A store manager wishes to see if the number of absences of her employees is the same for each weekday. She selected a random week and finds the following number of absences.

Day	Mon	Tues	Weds	Thurs	Fri
Absences	13	10	16	22	24

At $\alpha = 0.05$, is there a difference in the number of absences for each day of the week?

12. Ages of Head Start Program Students The Head Start Program provides a wide range of services to low-income children up to the age of 5 years and their families. Its goals are to provide services to improve social and learning skills and to improve health and nutrition status so that the participants can begin school on an equal footing with their more advantaged peers. The distribution of ages for participating children is as follows: 4% five-year-olds, 52% four-year-olds, 34% three-year-olds, and 10% under 3 years. When the program was assessed in a particular region, it was found that of the 200 randomly selected participants, 20 were 5 years old, 120 were 4 years old, 40 were 3 years old, and 20 were under 3 years. Is there sufficient evidence at $\alpha =$ 0.05 that the proportions differ from the program's? Use the *P*-value method.

*Source: New York Times Almanac/*www.fedstats.dhhs.gov

13. Firearms Deaths According to the National Safety Council, 10% of the annual deaths from firearms were victims from birth through 19 years of age. Half of the deaths from firearms were victims aged 20 through 44 years, and 40% of victims were aged 45 years and over. A random sample of 100 deaths by firearms in a particular state indicated the following: 13 were victims from birth through 19 years, 62 were aged 20 through 44 years, and the rest were 45 years old and older. At the 0.05 level of significance, are the results different from those cited by the National Safety Council?

Source: World Almanac.

14. College Degree Recipients A survey of 800 randomly selected recent degree recipients found that 155 received associate degrees; 450, bachelor degrees; 20, first professional degrees; 160, master degrees; and 15, doctorates. Is there sufficient evidence to conclude that at least one of the proportions differs from a report which stated that 23.3% were associate degrees; 51.1%, bachelor degrees; 3%, first professional degrees; 20.6%, master degrees; and 2%, doctorates? Use $\alpha = 0.05$.

Source: New York Times Almanac.

15. Internet Users A survey was targeted at determining if educational attainment affected Internet use. Randomly selected shoppers at a busy mall were asked if they used the Internet and their highest level of education attained. The results are listed below. Is there sufficient evidence at the 0.05 level of significance that the proportion of Internet users differs for any of the groups?

Graduated college +	Attended college	Did not attend
44	41	40

Source: www.infoplease.com

16. Types of Automobiles Purchased In a recent year U.S. retail automobile sales were categorized as listed below.

luxury 16.0% large 4.6% midsize 39.8% small 39.6%

A random sample of 150 recent purchases indicated the following results: 25 were luxury models, 12 were large cars, 60 were midsize, and 53 were small. At the 0.10 level of significance, is there sufficient evidence to conclude that the proportions of each type of car purchased differed from the report?

Source: World Almanac.

17. Paying for Prescriptions A medical researcher wishes to determine if the way people pay for their medical prescriptions is distributed as follows: 60% personal funds, 25% insurance, 15% Medicare. A random sample of 50 people found that 32 paid with their own money, 10 paid using insurance, and 8 paid using Medicare. At $\alpha = 0.05$, is the assumption correct? Use the *P*-value method. What would be an implication of the results?

Source: U.S. Health Care Financing.

18. Education Level and Health Insurance A researcher wishes to see if the number of randomly selected adults who do not have health insurance is equally distributed among three categories (less than 12 years of education, 12 years of education, more than 12 years of education). A sample of 60 adults who do not have health insurance is selected, and the results are shown. At $\alpha = 0.05$, can it be concluded that the frequencies are not equal? Use the *P*-value method. If the null hypothesis is rejected, give a possible reason for this.

Category	Less than 12 years	12 years	More than 12 years
Frequency	29	20	11

Source: U.S. Census Bureau.

Extending the Concepts

19. Tossing Coins Three coins are tossed 72 times, and the number of heads is shown. At $\alpha = 0.05$, test the null hypothesis that the coins are balanced and randomly tossed. (*Hint:* Use the binomial distribution.)

No. of heads	0	1	2	3
Frequency	3	10	17	42

20. State Lottery Numbers Select a three-digit state lottery number over a period of 50 days. Count the number of times each digit, 0 through 9, occurs. Test the claim, at $\alpha = 0.05$, that the digits occur at random.

Technology Step by Step

TI-84 Plus
Step by Step

Goodness-of-Fit Test

Example TI11–1

At the 5% significance level, test the claim that there is no preference in the selection of fruit soda flavors for the data.

Frequency	Cherry	Strawberry	Orange	Lime	Grape
Observed	32	28	16	14	10
Expected	20	20	20	20	20

To calculate the test statistic:

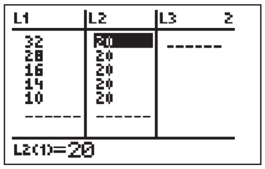
1. Enter the observed frequencies in L_1 and the expected frequencies in L_2.
2. Press **2nd [QUIT]** to return to the home screen.
3. Press **2nd [LIST],** move the cursor to MATH, and press **5** for sum(.
4. Type **$(L_1 - L_2)^2/L_2)$,** then press **ENTER.**

To calculate the *P*-value:

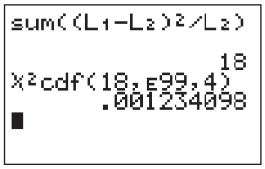
Press **2nd [DISTR],** then press **8** to get χ^2cdf(.
For this *P*-value, the χ^2cdf(command has the form χ^2cdf(test statistic, ∞, degrees of freedom).
Use E99 for ∞. Type **2nd [EE]** to get the small E.

Note: On newer TI84s, the χ^2cdf will come up as a menu. **Enter** the test statistic as the lower values, E99 as the upper, and 4 as df. After highlighting **PASTE** and pressing **ENTER** twice, the result will be the same as that shown.

For this example use χ^2cdf(18, E99,4):



```
L1      |L2      |L3      2
  32    |20      | ------
  28    |20      |
  16    |20      |
  14    |20      |
  10    |20      |
 ------ | ------ |
L2(1)=20
```



```
sum((L1-L2)²/L2)
                  18
χ²cdf(18,E99,4)
          .001234098
```

Since *P*-value = 0.001234098 < 0.05 = significance level, reject H_0 and conclude H_1. Therefore, there is enough evidence to reject the claim that consumers show no preference for soda flavors.

Note: On the newer TI-84 calculator, there is a goodness-of-fit test. Put the observed values in L1 and the expected values in L2 and then in TESTS-ALPHA D, enter df, and calculate.

EXCEL
Step by Step

Chi-Square Goodness-of-Fit Test

Excel does not have a procedure to conduct the goodness-of-fit test. However, you may conduct this test using the MegaStat Add-in available in your online resources. If you have not installed this add-in, do so, following the instructions from the Chapter 1 Excel Step by Step.

Example XL11–1

Test the claim that there is no preference for soda flavor. Use a significance level of $\alpha = 0.05$. The table of frequencies is shown below.

Frequency	**Cherry**	**Strawberry**	**Orange**	**Lime**	**Grape**
Observed	32	28	16	14	10
Expected	20	20	20	20	20

1. Enter the observed frequencies in row 1 (cells: A1 to E1) of a new worksheet.
2. Enter the expected frequencies in row 2 (cells: A2 to E2).
3. From the toolbar, select Add-Ins, **MegaStat>Chi-Square/Crosstab>Goodness of Fit Test.** *Note:* You may need to open MegaStat from the MegaStat.xls file on your computer's hard drive.
4. In the dialog box, type **A1:E1** for the Observed values and **A2:E2** for the Expected values. Then click [OK].

```
Goodness-of-Fit Test

Observed    Expected      O – E    (O – E)²/E    % of chisq
      32      20.000     12.000         7.200         40.00
      28      20.000      8.000         3.200         17.78
      16      20.000     –4.000         0.800          4.44
      14      20.000     –6.000         1.800         10.00
      10      20.000    –10.000         5.000         27.78
     100     100.000      0.000        18.000        100.00

   18.00    chi-square
       4    df
  0.0012    P-value
```

Since the *P*-value is less than the significance level, the null hypothesis is rejected and thus the claim of no preference is rejected.

Chi-Square Test for Normality

Example XL11–2

This example refers to Example 11–4. At the 5% significance level, determine if the variable is normally distributed.

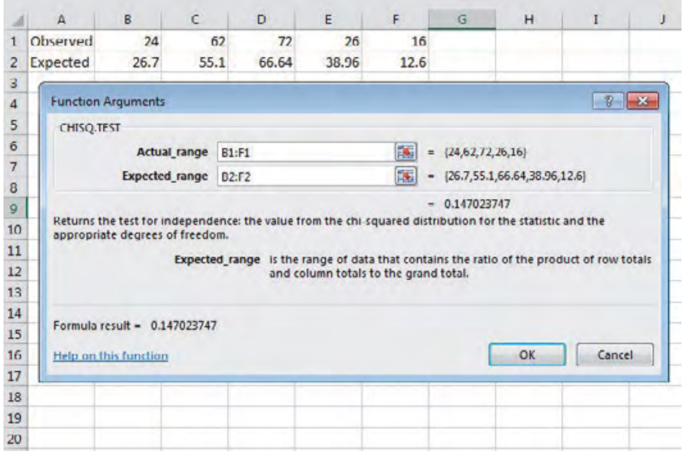
Start with the table of observed and expected values:

Observed	24	62	72	26	16
Expected	26.7	55.1	66.64	38.96	12.6

1. Enter the Observed values in row 1 of a new worksheet.
2. Enter the Expected values in row 2.

Note: You may include labels for the observed and expected values in cells A1 and A2, respectively.

3. Select the Insert Function Icon from the Toolbar.
4. In the Insert Function dialog box, select the Statistical category and the CHISQ.TEST function.
5. Type **B1:F1** for the Actual Range and **B2:F2** for the Expected Range, then click [OK].



	A	B	C	D	E	F
1	Observed	24	62	72	26	16
2	Expected	26.7	55.1	66.64	38.96	12.6

Function Arguments

CHISQ.TEST

Actual_range B1:F1 = {24,62,72,26,16}

Expected_range D2:F2 = {26.7,55.1,66.64,38.96,12.6}

= 0.147023747

Returns the test for independence: the value from the chi-squared distribution for the statistic and the appropriate degrees of freedom.

Expected_range is the range of data that contains the ratio of the product of row totals and column totals to the grand total.

Formula result = 0.147023747

The *P*-value of 0.1470 is greater than the significance level of 0.05. So we do not reject the null hypothesis. Thus, the distribution of the variable is approximately normal.

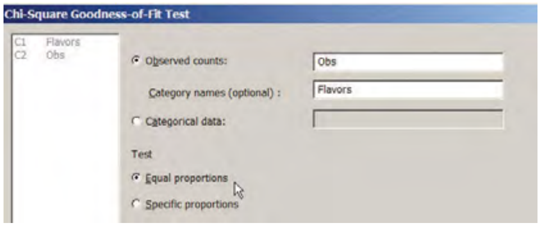
MINITAB
Step by Step

Chi-Square Goodness of Fit

Is there a preference for flavor of soda?

Frequency	Cherry	Strawberry	Orange	Lime	Grape
Observed	32	28	16	14	10
Expected	20	20	20	20	20

1. Enter the Observed Frequencies.
   a) Optional: type in the flavors in C1. Name the column Flavors.
   b) Type in the Observed counts in C2. Name the column Obs.
2. Select **Stat>Tables>Chi-Square Square Goodness-of-Fit Test (One Variable)**.
   a) Check the ratio button for Observed Counts then select C2 Obs.
   b) In the dialog box for Categorical names, select C1 Flavors (optional).
   c) Select the option for Equal proportions, then click [OK].



Chi-Square Goodness-of-Fit Test

C1 Flavors
C2 Obs

● Observed counts: Obs

Category names (optional) : Flavors

○ Categorical data:

Test

● Equal proportions

○ Specific proportions

The session window will contain the results as shown:

```
Chi-Square Goodness-of-Fit Test for Observed Counts in Variable: Obs
Using category names in Flavors
                               Test                 Contribution
Category       Observed   Proportion    Expected      to Chi-Sq
Cherry               32          0.2          20            7.2
Strawberry           28          0.2          20            3.2
Orange               16          0.2          20            0.8
Lime                 14          0.2          20            1.8
Grape                10          0.2          20            5.0
   N    DF    Chi-Sq    P-Value
 100     4        18      0.001
```

11–2 Tests Using Contingency Tables

When data can be tabulated in table form in terms of frequencies, several types of hypotheses can be tested by using the chi-square test.

Two such tests are the independence of variables test and the homogeneity of proportions test. The test of independence of variables is used to determine whether two variables are independent of or related to each other when a single sample is selected. The test of homogeneity of proportions is used to determine whether the proportions for a variable are equal when several samples are selected from different populations. Both tests use the chi-square distribution and a contingency table, and the test value is found in the same way. The independence test will be explained first.

OBJECTIVE ②

Test two variables for independence, using chi-square.

Test for Independence

The chi-square **independence test** is used to test whether two variables are independent of each other.

Formula for the Chi-Square Independence Test

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

with degrees of freedom equal to the (number of rows minus 1)(number of columns minus 1) and where

O = observed frequency
E = expected frequency

Assumptions for Chi-Square Independence Test

1. The data are obtained from a random sample.
2. The expected value in each cell must be 5 or more. If the expected values are not 5 or more, combine categories.

In this book, the assumptions will be stated in the exercises; however, when encountering statistics in other situations, you must check to see that these assumptions have been met before proceeding.

The null hypotheses for the chi-square independence test are generally, with some variations, stated as follows:

H_0: The variables are independent of each other.

H_1: The variables are dependent upon each other.

The data for the two variables are placed in a **contingency table.** One variable is called the *row variable,* and the other variable is called the *column variable.* The table is called an $R \times C$ *table,* where R is the number of rows and C is the number of columns. (Remember, rows go across or horizontally, and columns go up and down or vertically.) For example, a 2×3 contingency table would look like this.

	Column 1	Column 2	Column 3
Row 1	$C_{1,2}$	$C_{1,2}$	$C_{1,3}$
Row 2	$C_{2,1}$	$C_{2,2}$	$C_{2,3}$

Each value in the table is called a **cell value.** For example, the cell value $C_{2,3}$ means that it is in the second row (2) and third column (3).

The observed values are obtained from the sample data. (That is, they are given in the problem.) The expected values are computed from the observed values, and they are based on the assumption that the two variables are independent.

The formula for computing the expected values for each cell is

$$\text{Expected value} = \frac{(\text{row sum})(\text{column sum})}{\text{grand total}}$$

As with the chi-square goodness-of-fit test, if there is little difference between the observed values and the expected values, then the value of the test statistic will be small and the null hypothesis will not be rejected. Hence, the variables are independent of each other.

However, if there are large differences in the observed values and the expected values, then the test statistic will be large and the null hypothesis will be rejected. In this case, there is enough evidence to say that the variables are dependent on or related to each other. This test is always right-tailed.

The degrees of freedom for the chi-square test of independence are based on the size of the contingency table. The formula for the degrees of freedom is $(R - 1)(C - 1)$, that is, (rows − 1)(columns − 1). For a 2×3 table, the degrees of freedom would be $(2 - 1)(3 - 1) = 1(2) = 2$.

As an example, suppose a new postoperative procedure is administered to a number of patients in a large hospital. The researcher can ask the question, Do the doctors feel differently about this procedure from the nurses, or do they feel basically the same way? Note that the question is not whether they prefer the procedure but whether there is a difference of opinion between the two groups.

To answer this question, a researcher selects a sample of nurses and doctors and tabulates the data in table form, as shown.

Group	Prefer new procedure	Prefer old procedure	No preference
Nurses	100	80	20
Doctors	50	120	30

As the survey indicates, 100 nurses prefer the new procedure, 80 prefer the old procedure, and 20 have no preference; 50 doctors prefer the new procedure, 120 like the old procedure, and 30 have no preference. Since the main question is whether there is a difference in opinion, the null hypothesis is stated as follows:

H_0: The opinion about the procedure is *independent* of the profession.

The alternative hypothesis is stated as follows:

H_1: The opinion about the procedure is *dependent* on the profession.

If the null hypothesis is not rejected, the test means that both professions feel basically the same way about the procedure and that the differences are due to chance. If the null hypothesis is rejected, the test means that one group feels differently about the procedure from the other. Remember that rejection does *not* mean that one group favors the procedure and the other does not. Perhaps both groups favor it or both dislike it, but in different proportions.

Recall that the degrees of freedom for any contingency table are (rows − 1) times (columns − 1); that is, d.f. $= (R - 1)(C - 1)$. In this case, $(2 - 1)(3 - 1) = (1)(2) = 2$. The reason for this formula for d.f. is that all the expected values except one are free to vary in each row and in each column. The critical value for $\alpha = 0.05$ from Table G is 5.991.

To test the null hypothesis by using the chi-square independence test, you must compute the expected frequencies, assuming that the null hypothesis is true. These frequencies are computed by using the observed frequencies given in the table.

Interesting Facts

You're never too old—or too young—to be your best. George Foreman won the world heavyweight boxing championship at age 46. William Pitt was 24 when he became prime minister of Great Britain. Benjamin Franklin was a newspaper columnist at age 16 and a framer of the Constitution when he was 81.

Using the previous table, you can compute the expected frequencies for each block (or cell), as shown next.

1. Find the sum of each row and each column, and find the grand total, as shown.

Group	Prefer new procedure	Prefer old procedure	No preference	Total
Nurses	100	80	20	Row 1 sum 200
Doctors	+50	+120	+30	Row 2 sum 200
Total	150	200	50	400
	Column 1 sum	Column 2 sum	Column 3 sum	Grand total

2. For each cell, multiply the corresponding row sum by the column sum and divide by the grand total, to get the expected value:

$$\text{Expected value} = \frac{\text{row sum} \times \text{column sum}}{\text{grand total}}$$

For example, for $C_{1,2}$, the expected value, denoted by $E_{1,2}$, is (refer to the previous tables)

$$E_{1,2} = \frac{(200)(200)}{400} = 100$$

For each cell, the expected values are computed as follows:

$$E_{1,1} = \frac{(200)(150)}{400} = 75 \qquad E_{1,2} = \frac{(200)(200)}{400} = 100 \qquad E_{1,3} = \frac{(200)(50)}{400} = 25$$

$$E_{2,1} = \frac{(200)(150)}{400} = 75 \qquad E_{2,2} = \frac{(200)(200)}{400} = 100 \qquad E_{2,3} = \frac{(200)(50)}{400} = 25$$

The expected values can now be placed in the corresponding cells along with the observed values, as shown.

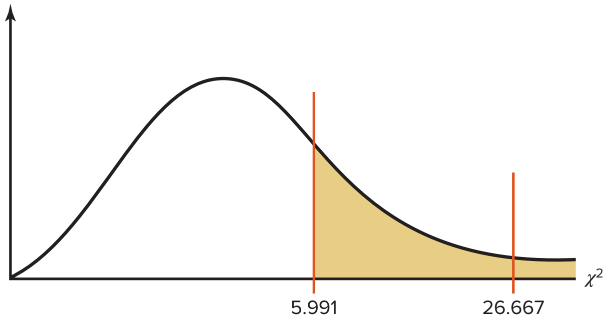
Group	Prefer new procedure	Prefer old procedure	No preference	Total
Nurses	100 (75)	80 (100)	20 (25)	200
Doctors	50 (75)	120 (100)	30 (25)	200
Total	150	200	50	400

The rationale for the computation of the expected frequencies for a contingency table uses proportions. For $C_{1,1}$ a total of 150 out of 400 people prefer the new procedure. And since there are 200 nurses, you would expect, if the null hypothesis were true, (150/400)(200), or 75, of the nurses to be in favor of the new procedure.

The test value can now be computed by using the formula

$$\begin{aligned}\chi^2 &= \sum \frac{(O - E)^2}{E} \\ &= \frac{(100 - 75)^2}{75} + \frac{(80 - 100)^2}{100} + \frac{(20 - 25)^2}{25} + \frac{(50 - 75)^2}{75} \\ &\quad + \frac{(120 - 100)^2}{100} + \frac{(30 - 25)^2}{25} \\ &= 26.667\end{aligned}$$

FIGURE 11–6
Critical and Test Values for the Postoperative Procedures Example



Next make the decision. Since the test value 26.667 is larger than the critical value 5.991, the decision is to reject the null hypothesis. See Figure 11–6.

The conclusion is that there is enough evidence to support the claim that opinion is related to (dependent on) profession, that is, that the doctors and nurses differ in their opinions about the procedure.

The *P*-values can be found as previously shown. In this case, if we look across row 2 (d.f. = 2), we see that the chi-square value of 10.597 corresponds to the α value of 0.005. Since our test value is 26.667 and is larger than 10.597, the *P*-value < 0.005, and the null hypothesis can be rejected.

The Procedure Table for performing the chi-square independence test is given next.

Procedure Table

The Chi-Square Independence Test

Step 1 State the hypotheses and identify the claim.

Step 2 Find the critical value for the right tail. Use Table G.

Step 3 Compute the test value. To compute the test value, first find the expected values. For each cell of the contingency table, use the formula

$$E = \frac{\text{(row sum)(column sum)}}{\text{grand total}}$$

to get the expected value. To find the test value, use the formula

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Step 4 Make the decision.

Step 5 Summarize the results.

Examples 11–5 and 11–6 illustrate the procedure for the chi-square test of independence.

EXAMPLE 11–5 Hospitals and Infections

A researcher wishes to see if there is a relationship between the hospital and the number of patient infections. A random sample of 3 hospitals was selected, and the number of infections for a specific year has been reported. The data are shown next.

Hospital	Surgical site infections	Pneumonia infections	Bloodstream infections	Total
A	41	27	51	119
B	36	3	40	79
C	169	106	109	384
Total	246	136	200	582

Source: Pennsylvania Health Care Cost Containment Council.

At $\alpha = 0.05$, can it be concluded that the number of infections is related to the hospital where they occurred?

Interesting Fact

There is enough water in the Great Lakes to cover the entire continental United States to a depth of $9\frac{1}{2}$ feet.

SOLUTION

Step 1 State the hypothesis and identify the claim.

H_0: The number of infections is independent of the hospital.
H_1: The number of infections is dependent on the hospital (claim).

Step 2 Find the critical value. The critical value using Table G at $\alpha = 0.05$ with $(3 - 1)(3 - 1) = (2)(2) = 4$ degrees of freedom is 9.488.

Step 3 Compute the test value. First find the expected values.

$$E_{1,1} = \frac{(119)(246)}{582} = 50.30 \qquad E_{1,2} = \frac{(119)(136)}{582} = 27.81 \qquad E_{1,3} = \frac{(119)(200)}{582} = 40.89$$

$$E_{2,1} = \frac{(79)(246)}{582} = 33.39 \qquad E_{2,2} = \frac{(79)(136)}{582} = 18.46 \qquad E_{2,3} = \frac{(79)(200)}{582} = 27.15$$

$$E_{3,1} = \frac{(384)(246)}{582} = 162.31 \qquad E_{3,2} = \frac{(384)(136)}{582} = 89.73 \qquad E_{3,3} = \frac{(384)(200)}{582} = 131.96$$

The completed table is shown.

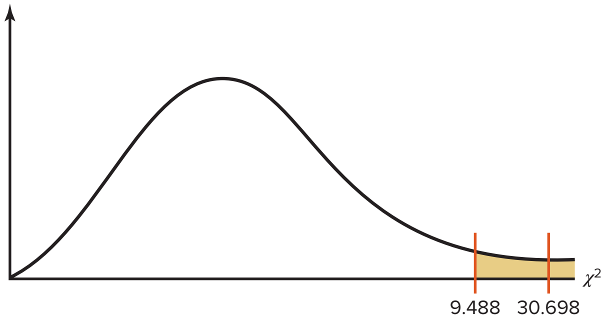
Hospital	Surgical site infections	Pneumonia infections	Bloodstream infections	Total
A	41 (50.30)	27 (27.81)	51 (40.89)	119
B	36 (33.39)	3 (18.46)	40 (27.15)	79
C	169 (162.31)	106 (89.73)	109 (131.96)	384
Total	246	136	200	582

Then substitute in the formula and evaluate to find the test statistic value.

$$\begin{aligned}
\chi^2 &= \sum \frac{(O - E)^2}{E} \\
&= \frac{(41 - 50.30)^2}{50.30} + \frac{(27 - 27.81)^2}{27.81} + \frac{(51 - 40.89)^2}{40.89} \\
&\quad + \frac{(36 - 33.39)^2}{33.39} + \frac{(3 - 18.46)^2}{18.46} + \frac{(40 - 27.15)^2}{27.15} \\
&\quad + \frac{(169 - 162.31)^2}{162.31} + \frac{(106 - 89.73)^2}{89.73} + \frac{(109 - 131.96)^2}{131.96} \\
&= 1.719 + 0.024 + 2.500 + 0.204 + 12.948 + 6.082 \\
&\quad + 0.276 + 2.950 + 3.995 \\
&= 30.698
\end{aligned}$$

Step 4 Make the decision. The decision is to reject the null hypothesis since $30.698 > 9.488$. That is, the test value lies in the critical region, as shown in Figure 11–7.

FIGURE 11–7
Critical and Test Values for Example 11–5



Step 5 Summarize the results. There is enough evidence to support the claim that the number of infections is related to the hospital where they occurred.

EXAMPLE 11–6 Mental Illness

A researcher wishes to see if there is a difference between men and women in the number of cases of mental disorders. She selects a sample of 30 males and 24 females and classifies them according to their mental disorder. The results are shown.

Gender	Anxiety	Depression	Schizophrenia	Total
Male	8	12	10	30
Female	12	9	3	24
Total	20	21	13	54

At $\alpha = 0.10$, can the researcher conclude that there is a difference in the types of disorders?

SOLUTION

Step 1 State the hypotheses and identify the claim.

H_0: The type of mental disorder is independent of the gender of the person.
H_1: The type of mental disorder is related to the gender of the person (claim).

Step 2 Find the critical value. The critical value is 4.605 since the degrees of freedom are $(2 - 1)(3 - 1) = 2$.

Step 3 Compute the test value. First compute the expected values.

$$E_{1,1} = \frac{(30)(20)}{54} = 11.11 \qquad E_{1,2} = \frac{(30)(21)}{54} = 11.67 \qquad E_{1,3} = \frac{(30)(13)}{54} = 7.22$$

$$E_{2,1} = \frac{(24)(20)}{54} = 8.89 \qquad E_{2,2} = \frac{(24)(21)}{54} = 9.33 \qquad E_{2,3} = \frac{(24)(13)}{54} = 5.78$$

The completed table is shown.

Gender	Anxiety	Depression	Schizophrenia	Total
Male	8 (11.11)	12 (11.67)	10 (7.22)	30
Female	12 (8.89)	9 (9.33)	3 (5.78)	24
Total	20	21	13	54

The test value is

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

$$= \frac{(8 - 11.11)^2}{11.11} + \frac{(12 - 11.67)^2}{11.67} + \frac{(10 - 7.22)^2}{7.22}$$

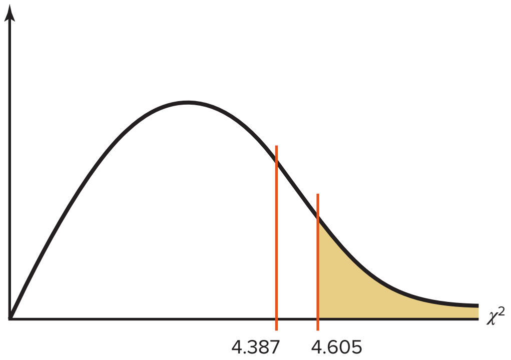
$$+ \frac{(12 - 8.89)^2}{8.89} + \frac{(9 - 9.33)^2}{9.33} + \frac{(3 - 5.78)^2}{5.78}$$

$$= 0.871 + 0.009 + 1.070 + 1.088 + 0.012 + 1.337 = 4.387$$

Step 4 Make the decision. The decision is to not reject the null hypothesis since $4.387 < 4.605$. See Figure 11–8.

FIGURE 11–8

Critical and Test Values for Example 11–6



Step 5 Summarize the results. There is not enough evidence to support the claim that the mental disorder is related to the gender of the individual.

In this case, the *P*-value is between 0.10 and 0.90. The TI-84 gives a *P*-value of 0.112. Again, this supports the decision and summary as stated previously.

Test for Homogeneity of Proportions

OBJECTIVE 3

Test proportions for homogeneity, using chi-square.

The second chi-square test that uses a contingency table is called the **homogeneity of proportions test.**

The *test of* **homogeneity of proportions** is used to test the claim that different populations have the same proportion of subjects who have a certain attitude or characteristic.

In this situation, samples are selected from several different populations, and the researcher is interested in determining whether the proportions of elements that have a common characteristic are the same for each population. The sample sizes are specified in advance, making either the row totals or column totals in the contingency table known before the samples are selected. For example, a researcher may select a sample of 50 freshmen, 50 sophomores, 50 juniors, and 50 seniors and then find the proportion of students who are smokers in each level. The researcher will then compare the proportions for each group to see if they are equal. The hypotheses in this case would be

H_0: $p_1 = p_2 = p_3 = p_4$

H_1: At least one proportion is different from the others.

If the researcher does not reject the null hypothesis, it can be assumed that the proportions are equal and the differences in them are due to chance. Hence, the proportion of

Interesting Facts

Water is the most critical nutrient in your body. It is needed for just about everything that happens. Water is lost fast: 2 cups daily are lost just exhaling, 10 cups through normal waste and body cooling, and 1 to 2 quarts per hour running, biking, or working out.

students who smoke is the same for grade levels freshmen through senior. When the null hypothesis is rejected, it can be assumed that the proportions are not all equal.

The assumptions for the test of homogeneity of proportions are the same as the assumptions for the chi-square test of independence. The procedure for this test is the same as the procedure for the chi-square test of independence.

EXAMPLE 11–7 Happiness and Income

A psychologist randomly selected 100 people from each of four income groups and asked them if they were "very happy." For people who made less than \$30,000, 24% responded yes. For people who made \$30,000 to \$74,999, 33% responded yes. For people who made \$75,000 to \$90,999, 38% responded yes, and for people who made \$100,000 or more, 49% responded yes. At $\alpha = 0.05$, test the claim that there is no difference in the proportion of people in each economic group who were very happy.

SOLUTION

It is necessary to make a table showing the number of people in each group who responded yes and the number of people in each group who responded no.

For group 1, 24% of the people responded yes, so 24% of 100 = 0.24(100) = 24 responded yes and 100 − 24 = 76 responded no.

For group 2, 33% of the people responded yes, so 33% of 100 = 0.33(100) = 33 responded yes and 100 − 33 = 67 responded no.

For group 3, 38% of the people responded yes, so 38% of 100 = 0.38(100) = 38 people responded yes and 100 − 38 = 62 people responded no.

For group 4, 49% of the people responded yes, so 49% of 100 = 0.49(100) = 49 responded yes, and 100 − 49 = 51 people responded no.

Tabulate the data in a table, and find the sums of the rows and columns as shown.

Household income	Less than \$30,000	\$30,000–\$74,999	\$75,000–\$99,999	\$100,000 or more	Total
Yes	24	33	38	49	144
No	76	67	62	51	256
	100	100	100	100	400

Source: Based on information from Princeton Survey Research Associates International.

Step 1 State the hypotheses and identify the claim.

H_0: $p_1 = p_2 = p_3 = p_4$ (claim)

H_1: At least one proportion differs from the others.

Step 2 Find the critical value. The formula for the degrees of freedom is the same as before: $(R - 1)(C - 1) = (2 - 1)(4 - 1) = 1(3) = 3$. The critical value is 7.815.

Step 3 Compute the test value. Since we want to test the claim that the proportions are equal, we use the expected value as $\frac{1}{4} \cdot 400 = 100$. First compute the expected values as shown previously.

$$E_{1,1} = \frac{(144)(100)}{400} = 36 \qquad E_{1,2} = \frac{(144)(100)}{400} = 36 \qquad E_{1,3} = \frac{(144)(100)}{400} = 36 \qquad E_{1,4} = \frac{(144)(100)}{400} = 36$$

$$E_{2,1} = \frac{(256)(100)}{400} = 64 \qquad E_{2,2} = \frac{(256)(100)}{400} = 64 \qquad E_{2,3} = \frac{(256)(100)}{400} = 64 \qquad E_{2,4} = \frac{(256)(100)}{400} = 64$$

The completed table is shown.

Household income	Less than $30,000	$30,000–$74,999	$75,000–$99,999	$100,000 or more	Total
Yes	24 (36)	33 (36)	38 (36)	49 (36)	144
No	76 (64)	67 (64)	62 (64)	51 (64)	256
	100	100	100	100	400

Next calculate the test value.

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

$$= \frac{(24-36)^2}{36} + \frac{(33-36)^2}{36} + \frac{(38-36)^2}{36} + \frac{(49-36)^2}{36}$$

$$+ \frac{(76-64)^2}{64} + \frac{(67-64)^2}{64} + \frac{(62-64)^2}{64} + \frac{(51-64)^2}{64}$$

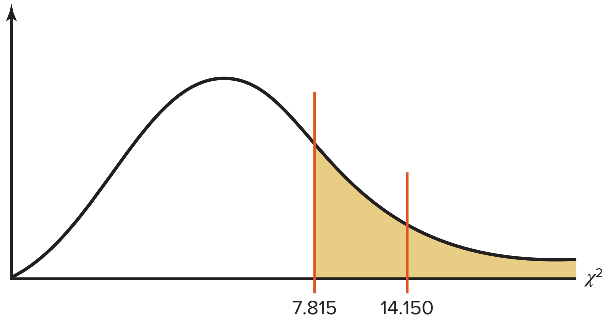
$$= 4.000 + 0.250 + 0.111 + 4.694 + 2.250 + 0.141 + 0.063 + 2.641$$

$$= 14.150$$

Step 4 Make the decision. Reject the null hypothesis since 14.150 > 7.815. See Figure 11–9.

FIGURE 11–9

Critical and Test Values for Example 11–7



Step 5 Summarize the results. There is enough evidence to reject the claim that there is no difference in the proportions. Hence, the incomes seem to make a difference in the proportions.

When the degrees of freedom for a contingency table are equal to 1—that is, the table is a 2 × 2 table—some statisticians suggest using the *Yates correction for continuity*. The formula for the test is then

$$\chi^2 = \sum \frac{(|O-E| - 0.5)^2}{E}$$

Since the chi-square test is already conservative, most statisticians agree that the Yates correction is not necessary. (See Exercise 33 in Extending the Concepts.)

Applying the Concepts 11–2

Satellite Dishes in Restricted Areas

The Senate is expected to vote on a bill to allow the installation of satellite dishes of any size in deed-restricted areas. The House had passed a similar bill. An opinion poll was taken to see whether how a person felt about satellite dish restrictions was related to his or her age. A chi-square test was run, creating the following computer-generated information.

Degrees of freedom d.f. = 6
Test statistic $\chi^2 = 61.25$
Critical value C.V. = 12.6
P-value = 0.00
Significance level = 0.05

	18–29	30–49	50–64	65 and up
For	96 (79.5)	96 (79.5)	90 (79.5)	36 (79.5)
Against	201 (204.75)	189 (204.75)	195 (204.75)	234 (204.75)
Don't know	3 (15.75)	15 (15.75)	15 (15.75)	30 (15.75)

1. Which number from the output is compared to the significance level to check if the null hypothesis should be rejected?
2. Which number from the output gives the probability of a type I error that is calculated from your sample data?
3. Was a right-, left-, or two-tailed test run? Why?
4. Can you tell how many rows and columns there were by looking at the degrees of freedom?
5. Does increasing the sample size change the degrees of freedom?
6. What are your conclusions? Look at the observed and expected frequencies in the table to draw some of your own specific conclusions about response and age.
7. What would your conclusions be if the level of significance were initially set at 0.10?
8. Does chi-square tell you which cell's observed and expected frequencies are significantly different?

See page 643 for the answers.

Exercises 11–2

1. How is the chi-square independence test similar to the goodness-of-fit test? How is it different?

2. How are the degrees of freedom computed for the independence test?

3. Generally, how would the null and alternative hypotheses be stated for the chi-square independence test?

4. What is the name of the table used in the independence test?

5. How are the expected values computed for each cell in the table?

6. How are the null and alternative hypotheses stated for the test of homogeneity of proportions?

For Exercises 7 through 31, perform the following steps.

a. State the hypotheses and identify the claim.
b. Find the critical value.
c. Compute the test value.
d. Make the decision.
e. Summarize the results.

Use the traditional method of hypothesis testing unless otherwise specified. Assume all assumptions are valid.

7. Living Arrangements A recent study of 100 individuals found the following living arrangement for men and women. The results are shown. Check the data for a dependent relationship at $\alpha = 0.05$.

	Spouse	Relative	Nonrelative	Alone
Men	57	8	25	10
Women	53	5	28	14

Source: Based on information from U.S. Census Bureau.

8. **Ethnicity and Movie Admissions** Are movie admissions related to ethnicity? A 2014 study indicated the following numbers of admissions (in thousands) for two different years. At the 0.05 level of significance, can it be concluded that movie attendance by year was dependent upon ethnicity?

	Caucasian	Hispanic	African American	Other
2013	724	335	174	107
2014	370	292	152	140

Source: MPAA Study.

9. **Pet Owners** A study of pet owners showed the following information concerning the ownership of the dogs and cats. At $\alpha = 0.10$, is there a relationship between the number of people in a family and dog or cat ownership?

	1 person	2 people	3 people	4 or more people
Dog	7	16	11	16
Cat	9	14	16	11

10. **Women in the Military** This table lists the numbers of officers and enlisted personnel for women in the military. At $\alpha = 0.05$, is there sufficient evidence to conclude that a relationship exists between rank and branch of the Armed Forces?

	Officers	Enlisted
Army	10,791	62,491
Navy	7,816	42,750
Marine Corps	932	9,525
Air Force	11,819	54,344

Source: New York Times Almanac.

11. **Violent Crimes** A record of violent crimes for a random sample of cities from the list of U.S. cities with the fewest crimes is shown below. At $\alpha = 0.05$, is there sufficient evidence to indicate a relationship between the city and the type of crime committed?

	Forcible rape	Robbery	Aggravated assault
Cary, NC	14	35	70
Amherst, NY	10	33	76
Simi Valley, CA	14	37	77
Norman, OK	47	36	53

Source: Time Almanac.

12. **Population and Age** Is the size of the population by age related to the state that it's in? Use $\alpha = 0.05$. (Population values are in thousands.)

	Under 5	5–17	18–24	25–44	45–64	65+
Pennsylvania	721	2140	1025	3515	2702	1899
Ohio	740	2104	1065	3359	2487	1501

Source: New York Times Almanac.

13. **Unemployment Time and Type of Industry** Is the length of unemployment related to the type of industry? A random sample of unemployed workers from three different sectors yielded the following data. At the 0.05 level of significance, are the two categories dependent?

	Less than 5 weeks	5–14 weeks	15–26 weeks
Transportation/ utilities	85	110	80
Information	48	57	45
Financial activities	83	111	114

Source: World Almanac.

14. **Congressional Representatives** Four states were randomly selected, and their members in the U.S. House of Representatives (114th Congress) are noted. At $\alpha = 0.10$, can it be concluded that there is a dependent relationship between the state and the political party affiliation of its representatives?

	California	Florida	Illinois	Texas
Democrat	39	10	10	11
Republican	14	17	8	25

15. **Student Majors at Colleges** The table shows the number of students (in thousands) participating in various programs at both two-year and four-year institutions. At $\alpha = 0.05$, can it be concluded that there is a relationship between program of study and type of institution?

	Two-year	Four-year
Agriculture and related sciences	36	52
Criminal justice	210	231
Foreign languages and literature	28	59
Mathematics and statistics	28	63

Source: Time Almanac.

16. **Organ Transplantation** Listed is information regarding organ transplantation for three different years. Based on these data, is there sufficient evidence at $\alpha = 0.01$ to conclude that a relationship exists between year and type of transplant?

Year	Heart	Kidney/pancreas	Lung
2003	2056	870	1085
2004	2016	880	1173
2005	2127	903	1408

Source: www.infoplease.com

17. Automobile Ownership A study was done on the type of automobiles owned by women and men. The data are shown. At $\alpha = 0.10$, is there a relationship between the type of automobile owned and the gender of the individual?

	Luxury	Large	Midsize	Small
Men	15	9	49	27
Women	9	6	62	14

18. Music CDs Sold Are the sales of CDs (in millions) by genre related to the year in which the sales occurred? Use the 0.05 level of significance.

	R&B	Country	Rock
2013	48	36	93
2014	36	30	85

Source: statista.com

19. Type of Medicine A medical researcher wishes to determine if the type of vitamin pills taken is related to the age of the person. The data are shown. At $\alpha = 0.10$, is the type of vitamin related to the age of the person taking the vitamin?

	Type		
Age	**Liquid**	**Tablet**	**Gummy**
20–39	5	16	6
40–59	10	25	12
60–over	19	6	5

20. Effectiveness of New Drug To test the effectiveness of a new drug, a researcher gives one group of randomly selected individuals the new drug and another group of randomly selected individuals a placebo. The results of the study are shown here. At $\alpha = 0.10$, can the researcher conclude that the drug results differ from those of the placebo? Use the *P*-value method.

Medication	Effective	Not effective
Drug	32	9
Placebo	12	18

21. Hospitals and Cesarean Delivery Rates The national Cesarean delivery rate for a recent year was 32.2% (number of live births performed by Cesarean section). A random sample of 100 birth records from three large hospitals showed the following results for type of birth. Test for homogeneity of proportions using $\alpha = 0.10$.

	Hospital A	Hospital B	Hospital C
Cesarean	44	28	39
Non-Cesarean	56	72	61

Source: World Almanac.

22. Foreign Language Speaking Dorms A local college recently made the news by offering foreign language–speaking dorm rooms to its students. When questioned at another school, 50 randomly selected students from each class responded as shown. At $\alpha = 0.05$, is there sufficient evidence to conclude that the proportions of students favoring foreign language–speaking dorms are not the same for each class?

	Freshmen	Sophomores	Juniors	Seniors
Yes (favor)	10	15	20	22
No	40	35	30	28

23. Youth Physical Fitness According to a recent survey, 64% of Americans between the ages of 6 and 17 years cannot pass a basic fitness test. A physical education instructor wishes to determine if the percentages of such students in different schools in his school district are the same. He administers a basic fitness test to 120 randomly selected students in each of four schools. The results are shown here. At $\alpha = 0.05$, test the claim that the proportions who pass the test are equal.

	Southside	West End	East Hills	Jefferson
Passed	49	38	46	34
Failed	71	82	74	86
Total	120	120	120	120

Source: The Harper's Index Book.

24. Participation in Market Research Survey An advertising firm has decided to ask 92 customers at each of three local shopping malls if they are willing to take part in a market research survey. According to previous studies, 38% of Americans refuse to take part in such surveys. The results are shown here. At $\alpha = 0.01$, test the claim that the proportions of those who are willing to participate are equal.

	Mall A	Mall B	Mall C
Will participate	52	45	36
Will not participate	40	47	56
Total	92	92	92

Source: The Harper's Index Book.

25. Athletic Status and Meat Preference A study was done using a sample of 60 college athletes and 60 college students who were not athletes. They were asked their meat preference. The data are shown. At $\alpha = 0.05$, test the claim that the preference proportions are the same.

	Pork	Beef	Poultry
Athletes	15	36	9
Nonathletes	17	28	15

26. Mothers Working Outside the Home According to a recent survey, 59% of Americans aged 8 to 17 years would prefer that their mothers work outside the home, regardless of what they do now. A school district psychologist decided to select three random samples of 60 students each in elementary, middle, and high school to see how the students in her district felt about the issue. At $\alpha = 0.10$, test the claim that the proportions of the students who prefer that their mothers have jobs outside the home are equal.

	Elementary	Middle	High
Prefer mothers work	29	38	51
Prefer mothers not work	31	22	9
Total	60	60	60

Source: Daniel Weiss, *100% American.*

27. **Volunteer Practices of Students** The Bureau of Labor Statistics reported information on volunteers by selected characteristics. They found that 24.4% of the population aged 16 to 24 years volunteers a median number of 36 hours per year. A survey of 75 randomly selected students in each age group revealed the following data on volunteer practices. At $\alpha = 0.05$, can it be concluded that the proportions of volunteers are the same for each group?

	Age				
	18	19	20	21	22
Yes (volunteer)	19	18	23	31	13
No	56	57	52	44	62

Source: Time Almanac.

28. **Fathers in the Delivery Room** On average, 79% of American fathers are in the delivery room when their children are born. A physician's assistant surveyed 300 randomly selected first-time fathers to determine if they had been in the delivery room when their children were born. The results are shown here. At $\alpha = 0.05$, is there enough evidence to reject the claim that the proportions of those who were in the delivery room at the time of birth are the same?

	Hospital A	Hospital B	Hospital C	Hospital D
Present	66	60	57	56
Not present	9	15	18	19
Total	75	75	75	75

Source: Daniel Weiss, *100% American.*

29. **Injuries on Monkey Bars** A children's playground equipment manufacturer read in a survey that 55% of all U.S. playground injuries occur on the monkey bars. The manufacturer wishes to investigate playground injuries in four different parts of the country to determine if the proportions of accidents on the monkey bars are equal. The results are shown here. At $\alpha = 0.05$, test the claim that the proportions are equal. Use the P-value method.

Accidents	North	South	East	West
On monkey bars	15	18	13	16
Not on monkey bars	15	12	17	14
Total	30	30	30	30

Source: Michael D. Shook and Robert L. Shook, *The Book of Odds.*

30. **Thanksgiving Travel** According to the American Automobile Association, 31 million Americans travel over the Thanksgiving holiday. To determine whether to stay open or not, a national restaurant chain surveyed 125 customers at each of four locations to see if they would be traveling over the holiday. The results are shown here. At $\alpha = 0.10$, test the claim that the proportions of Americans who will travel over the Thanksgiving holiday are equal. Use the P-value method.

	Location A	Location B	Location C	Location D
Will travel	37	52	46	49
Will not travel	88	73	79	76
Total	125	125	125	125

Source: Michael D. Shook and Robert L. Shook, *The Book of Odds.*

31. **Age and Drug Use** A study was done on people convicted of using illegal drugs and their ages. The data are shown. At $\alpha = 0.10$, test the claim that the proportions of the type of drug used are the same for the three age groups.

Age	Inhalants	Hallucinogens	Tranquilizers
12–17	16	9	5
18–25	22	30	8
26 and older	13	18	10

Extending the Concepts

32. For a 2×2 table, a, b, c, and d are the observed values for each cell, as shown.

a	b
c	d

The chi-square test value can be computed as

$$\chi^2 = \frac{n(ad - bc)^2}{(a + b)(a + c)(c + d)(b + d)}$$

SPEAKING OF STATISTICS Does Color Affect Your Appetite?

It has been suggested that color is related to appetite in humans. For example, if the walls in a restaurant are painted certain colors, it is thought that the customer will eat more food. A study was done at the University of Illinois and the University of Pennsylvania. When people were given six varieties of jellybeans mixed in a bowl or separated by color, they ate about twice as many from the bowl with the mixed jellybeans as from the bowls that were separated by color.

It is thought that when the jellybeans were mixed, people felt that it offered a greater variety of choices, and the variety of choices increased their appetites.

In this case one variable—color—is categorical, and the other variable—amount of jellybeans eaten—is numerical. Could a chi-square goodness-of-fit test be used here? If so, suggest how it could be set up.

© Siede Preis/Getty Images RF

where $n = a + b + c + d$. Using this formula, compute the χ^2 test value and then the formula $\Sigma(O - E)^2/E$, and compare the results. Use the following table.

12	15
9	23

33. For the contingency table shown in Exercise 32, compute the chi-square test value by using the Yates correction (page 630) for continuity.

34. When the chi-square test value is significant and there is a relationship between the variables, the strength of this relationship can be measured by using the *contingency coefficient.* The formula for the contingency coefficient is

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}}$$

where χ^2 is the test value and n is the sum of frequencies of the cells. The contingency coefficient will always be less than 1. Compute the contingency coefficient for Exercises 8 and 20.

Technology Step by Step

TI-84 Plus

Step by Step

Chi-Square Test for Independence

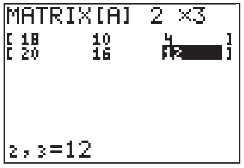
1. Press **2nd [X^{-1}]** for **MATRIX** and move the cursor to Edit; then press **ENTER.**
2. Enter the number of rows and columns. Then press **ENTER.**
3. Enter the values in the matrix as they appear in the contingency table.
4. Press **STAT** and move the cursor to TESTS. Press **C (ALPHA PRGM)** for χ^2-Test. Make sure the observed matrix is [A] and the expected matrix is [B].
5. Move the cursor to Calculate and press **ENTER.**

Example TI11–2

Test the claim of independence at $\alpha = 0.10$.

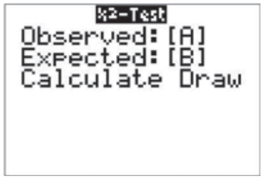
	Football	Baseball	Hockey
Male	18	10	4
Female	20	16	12

Input



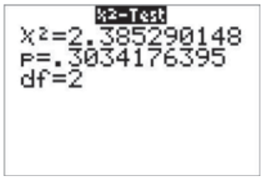
```
MATRIX[A] 2 ×3
[ 18      10      4  ]
[ 20      16      12 ]

2,3=12
```

Input



```
χ2-Test
Observed:[A]
Expected:[B]
Calculate Draw
```

Output



```
χ2-Test
χ²=2.385290148
p=.3034176395
df=2
```

The test value is 2.385290148. The *P*-value is 0.3034176395. The decision is to not reject the null hypothesis, since this value is greater than 0.10. You can find the expected values by pressing **MATRIX,** moving the cursor to [B], and pressing **ENTER** twice.

EXCEL
Step by Step

Tests Using Contingency Tables

Excel does not have a procedure to conduct tests using contingency tables without including the expected values. However, you may conduct such tests using the MegaStat Add-in available in your online resources. If you have not installed this add-in, do so, following the instructions from the Chapter 1 Excel Step by Step.

Example XL11–3

The table below shows the number of years of college a person has completed and the residence of the person.

Using a significance level $\alpha = 0.05$, determine whether the number of years of college a person has completed is related to residence.

1. Enter the location variable labels in column A, beginning at cell A2.
2. Enter the categories for the number of years of college in cells B1, C1, and D1, respectively.
3. Enter the observed values in the appropriate block (cell).
4. From the toolbar, select Add-Ins, **MegaStat>Chi-Square/Crosstab>Contingency Table.** *Note:* You may need to open MegaStat from the MegaStat.xls file on your computer's hard drive.
5. In the dialog box, type **A1:D4** for the Input range.
6. Check chi-square from the Output Options.
7. Click [OK].

Chi-Square Contingency Table Test for Independence

	None	4-year	Advanced	Total
Urban	15	12	8	35
Suburban	8	15	9	32
Rural	6	8	7	21
Total	29	35	24	88
	3.01	chi-square		
	4	df		
	.5569	*P*-value		

The results of the test indicate that at the 5% level of significance, there is not enough evidence to conclude that a person's location is dependent on number of years of college.

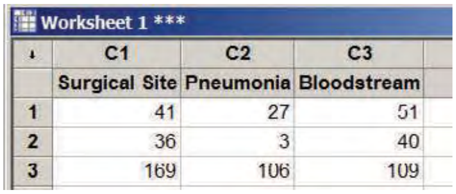
MINITAB
Step by Step

Chi-Square Test of Independence from Contingency Table

Example 11–5

Is there a relationship between the type of infection and the hospital?

1. Enter the Observed Frequencies for the type of infection in C1 Surgical Site, C2 Pneumonia, and C3 Bloodstream. Do not include labels or totals.



↓	C1	C2	C3
	Surgical Site	Pneumonia	Bloodstream
1	41	27	51
2	36	3	40
3	169	106	109

2. Select **Stat>Tables>Chi-Square Test for Association.**
   a) In the drop down menu select Summarized data in a two-way table.
   b) Click the text box for Columns containing the table and then double click each column name to include it in the list.
   c) Click on [Select], then [OK].

The results are displayed in the session window.

```
Chi-Square Test: Surgical Site, Pneumonia, Bloodstream
Expected counts are printed below observed counts
Chi-Square contributions are printed below expected counts

          Surgical
              Site    Pneumonia    Bloodstream    Total
    1           41           27             51      119
             50.30        27.81          40.89
             1.719        0.023          2.498

    2           36            3             40       79
             33.39        18.46          27.15
             0.204       12.948          6.084

    3          169          106            109      384
            162.31        89.73         131.96
             0.276        2.949          3.994

Total          246          136            200      582
Chi-Sq = 30.696, DF = 4, P-Value = 0.000
```

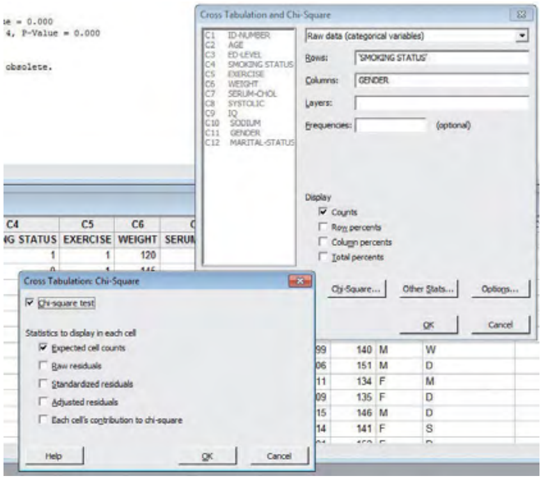
There is a relationship between infection type and hospital.

Construct a Contingency Table and Calculate the Chi-Square Test Statistic

In Chapter 4 we learned how to construct a contingency table by using gender and smoking status in the Data Bank file described in Appendix B. Are smoking status and gender related? Who is more likely to smoke, men or women?

1. Use **File>Open Worksheet** to open the Data Bank file. Remember, do *not* click the file icon.
2. Select **Stat>Tables>Cross Tabulation and Chi-Square.**
3. Make sure that Raw data (categorical variables) is selected in the drop down menu. Double-click Smoking Status for rows and Gender for columns.
4. The Display option for Counts should be checked.
5. Click [Chi-Square].
   a) Check Chi-Square test.
   b) Check Expected cell counts.
6. Click [OK] twice.

In the session window the contingency table and the chi-square analysis will be displayed.



Raw data (categorical variables)

Rows: 'SMOKING STATUS'

Columns: GENDER

Layers:

Frequencies: (optional)

Display

☑ Counts

☐ Row percents

☐ Column percents

☐ Total percents

☑ Chi-square test

Statistics to display in each cell

☑ Expected cell counts

☐ Raw residuals

☐ Standardized residuals

☐ Adjusted residuals

☐ Each cell's contribution to chi-square

```
Tabulated statistics: SMOKING STATUS, GENDER
Rows:  SMOKING      STATUS       Columns:  Gender

                F           M          All

 0             25          22           47
            23.50       23.50        47.00

 1             18          19           37
            18.50       18.50        37.00

 2              7           9           16
             8.00        8.00        16.00

 All           50          50          100
            50.00       50.00       100.00
 Cell Contents:               Count
                              Expected count

Pearson Chi-Square = 0.469, DF = 2, P-Value = 0.791
```

There is not enough evidence to conclude that smoking is related to gender.

Summary

- Three uses of the chi-square distribution were explained in this chapter. It can be used as a goodness-of-fit test to determine whether the frequencies of a distribution are the same as the hypothesized frequencies. For example, is the number of defective parts produced by a factory the same each day? This test is always a right-tailed test. (11–1)
- The test of independence is used to determine whether two variables are related or are independent. This test uses a contingency table and is always a right-tailed test. An example of its use is a test to determine if attitudes about trash recycling are dependent on whether residents live in urban or rural areas. (11–2)
- Finally, the homogeneity of proportions test is used to determine if several proportions are all equal when samples are selected from different populations. (11–2)

The chi-square distribution is also used for other types of statistical hypothesis tests, such as the Kruskal-Wallis test, which is explained in Chapter 13.

Important Terms

cell value 622

contingency table 622

expected frequency 608

chi-square goodness-of-fit test 608

homogeneity of proportions test 628

independence test 622

observed frequency 608

Important Formulas

Formula for the chi-square test for goodness of fit:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

with degrees of freedom equal to the (number of rows minus 1)(number of columns minus 1) and where

O = observed frequency

E = expected frequency

Formula for the chi-square independence and homogeneity of proportions tests:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

with degrees of freedom equal to (rows − 1) times (columns − 1). Formula for the expected value for each cell:

$$E = \frac{\textbf{(row sum)(column sum)}}{\textbf{grand total}}$$

Review Exercises

For Exercises 1 through 10, follow these steps.

a. State the hypotheses and identify the claim.
b. Find the critical value(s).
c. Compute the test value.
d. Make the decision.
e. Summarize the results.

Use the traditional method of hypothesis testing unless otherwise specified. Assume all assumptions have been met.

Section 11–1

1. **Traffic Accident Fatalities** A traffic safety report indicated that for the 21–24 year age group, 31.58% of traffic fatalities were victims who had used a seat belt. Victims who were not wearing a seat belt accounted for 59.83% of the deaths, and the status of the rest was unknown. A study of 120 randomly selected traffic fatalities in a particular region showed that for this age group, 35 of the victims had used a seat belt, 78 had not, and the status of the rest was unknown. At $\alpha = 0.05$, is there sufficient evidence that the proportions differ from those in the report?

Source: New York Times Almanac.

2. **Displaced Workers** The reasons that workers in the 25–54 year old category were displaced are listed.

Plant closed/moved	44.8%
Insufficient work	25.2%
Position eliminated	30%

A random sample of 180 displaced workers (in this age category) found that 40 lost their jobs due to their position being eliminated, 53 due to insufficient work, and the rest due to the company being closed or moving. At the 0.01 level of significance, are these proportions different from those from the U.S. Department of Labor?

Source: BLS-World Almanac.

3. **Gun Sale Denials** A police investigator read that the reasons why gun sales to applicants were denied were distributed as follows: criminal history of felonies, 75%; domestic violence conviction, 11%; and drug abuse, fugitive, etc., 14%. A random sample of applicants in a large study who were refused sales is obtained and is distributed as follows. At $\alpha = 0.10$, can it be concluded that the distribution is as stated? Do you think the results might be different in a rural area?

Reason	Criminal history	Domestic violence	Drug abuse, etc.
Number	120	42	38

Source: Based on FBI statistics.

4. **Type of Music Preferred** A store manager wishes to see if the employees of the store had a preference in the types of music played on the speaker system. A survey of 60 employees stated the type of music they preferred. The data are shown. At $\alpha = 0.05$, is there a preference?

Type	**Classical**	**Country**	**Pop**	**Rock**
Number	9	18	22	11

Section 11–2

5. **Pension Investments** A survey was conducted on how a lump-sum pension would be invested by randomly selected 45-year-olds and randomly selected 65-year-olds. The data are shown here. At $\alpha = 0.05$, is there a relationship between the age of the investor and the way the money would be invested?

	Large company stock funds	**Small company stock funds**	**International stock funds**	**CDs or money market funds**	**Bonds**
Age 45	20	10	10	15	45
Age 65	42	24	24	6	24

Source: USA TODAY.

6. **Tornadoes** According to records from the Storm Prediction Center, the following numbers of tornadoes occurred in the first quarter of each of years 2012–2015. Is there sufficient evidence to conclude that a relationship exists between the month and year in which the tornadoes occurred? Use $\alpha = 0.05$.

	2015	**2014**	**2013**	**2012**
January	26	4	87	97
February	2	41	46	63
March	13	25	18	225

Source: National Weather Service Storm Prediction Center.

7. **Employment of High School Females** A guidance counselor wishes to determine if the proportions of female high school students in his school district who have jobs are equal to the national average of 36%. He randomly surveys 80 female students, ages 16 through 18 years, to determine if they work. The results are shown. At $\alpha = 0.01$, test the claim that the proportions of female students who work are equal. Use the *P*-value method.

	16-year-olds	**17-year-olds**	**18-year-olds**
Work	45	31	38
Don't work	35	49	42
Total	80	80	80

Source: Michael D. Shook and Robert L. Shook, *The Book of Odds.*

8. **Risk of Injury** The risk of injury is higher for males compared to females (57% versus 43%). A hospital emergency room supervisor wishes to determine if the proportions of injuries to males in his hospital are the

STATISTICS TODAY

Statistics and Heredity—Revisited

Using probability, Mendel predicted the following:

	Smooth		Wrinkled	
	Yellow	**Green**	**Yellow**	**Green**
Expected	0.5625	0.1875	0.1875	0.0625

The observed results were these:

	Smooth		Wrinkled	
	Yellow	**Green**	**Yellow**	**Green**
Observed	0.5666	0.1942	0.1816	0.0556

Using chi-square tests on the data, Mendel found that his predictions were accurate in most cases (i.e., a good fit), thus supporting his theory. He reported many highly successful experiments. Mendel's genetic theory is simple but useful in predicting the results of hybridization.

A Fly in the Ointment

Although Mendel's theory is basically correct, an English statistician named R. A. Fisher examined Mendel's data some 50 years later. He found that the observed (actual) results agreed too closely with the expected (theoretical) results and concluded that the data had been falsified in some way. The results were too good to be true. Several explanations have been proposed, ranging from deliberate misinterpretation to an assistant's error, but no one can be sure how this happened.

same for each of 4 months. He randomly surveys 100 injuries treated in his ER for each month. The results are shown. At $\alpha = 0.05$, can he reject the claim that the proportions of injuries for males are equal for each of the four months?

	May	June	July	August
Male	51	47	58	63
Female	49	53	42	37
Total	100	100	100	100

Source: Michael D. Shook and Robert L. Shook, *The Book of Odds.*

9. **Living Arrangements** A study was done on the living arrangement of individuals aged 18 to 24 years to see if the proportions of people are different today than those of 10 years ago. The data are shown. At $\alpha = 0.05$, test the claim.

	Parents	Spouse	Non–family member	Other
10 years ago	47	38	5	10
Today	57	20	9	14

10. **Cardiovascular Procedures** Is the frequency of cardiovascular procedure related to gender? The following data were obtained for selected procedures for a recent year. At $\alpha = 0.10$, is there sufficient evidence to conclude a dependent relationship between gender and procedure?

	Coronary artery stent	Coronary artery bypass	Pacemaker
Men	425	320	198
Women	227	123	219

Source: New York Times Almanac.

Data Analysis

The Data Bank is located in Appendix B, or on the World Wide Web by following links from www.mhhe.com/math/stat/bluman

1. Select a random sample of 40 individuals from the Data Bank. Use the chi-square goodness-of-fit test to see if the marital status of individuals is equally distributed.
2. Use the chi-square test of independence to test the hypothesis that smoking is independent of gender. Use a random sample of at least 75 people.
3. Using the data from Data Set X in Appendix B, classify the data as 1–3, 4–6, 7–9, etc. Use the chi-square goodness-of-fit test to see if the number of times each ball is drawn is equally distributed.

Chapter Quiz

Determine whether each statement is true or false. If the statement is false, explain why.

1. The chi-square test of independence is always two-tailed.
2. The test values for the chi-square goodness-of-fit test and the independence test are computed by using the same formula.
3. When the null hypothesis is rejected in the goodness-of-fit test, it means there is close agreement between the observed and expected frequencies.

Select the best answer.

4. The values of the chi-square variable cannot be
   - *a.* Positive
   - *b.* 0
   - *c.* Negative
   - *d.* None of the above
5. The null hypothesis for the chi-square test of independence is that the variables are
   - *a.* Dependent
   - *b.* Independent
   - *c.* Related
   - *d.* Always 0
6. The degrees of freedom for the goodness-of-fit test are
   - *a.* 0
   - *b.* 1
   - *c.* Sample size − 1
   - *d.* Number of categories − 1

Complete the following statements with the best answer.

7. The degrees of freedom for a 4 × 3 contingency table are _______.
8. An important assumption for the chi-square test is that the observations must be _______.
9. The chi-square goodness-of-fit test is always _______-tailed.
10. In the chi-square independence test, the expected frequency for each class must always be _______.

For Exercises 11 through 19, follow these steps.

- *a.* State the hypotheses and identify the claim.
- *b.* Find the critical value.
- *c.* Compute the test value.
- *d.* Make the decision.
- *e.* Summarize the results.

Use the traditional method of hypothesis testing unless otherwise specified.

11. **Job Loss Reasons** A survey of why randomly selected people lost their jobs produced the following results. At $\alpha = 0.05$, test the claim that the number of responses is equally distributed. Do you think the results might be different if the study were done 10 years ago?

Reason	Company closing	Position abolished	Insufficient work
Number	26	18	28

Source: Based on information from U.S. Department of Labor.

12. **Consumption of Takeout Foods** A food service manager read that the place where people consumed takeout food is distributed as follows: home, 53%; car, 19%; work, 14%; other, 14%. A survey of 300 randomly selected individuals showed the following results. At $\alpha = 0.01$, can it be concluded that the distribution is as stated? Where would a fast-food restaurant want to target its advertisements?

Place	Home	Car	Work	Other
Number	142	57	51	50

Source: Beef Industry Council.

13. **Television Viewing** A survey of randomly selected people found that 62% of the respondents stated that they never watched the home shopping channels on cable television, 23% stated that they watched the channels rarely, 11% stated that they watched them occasionally, and 4% stated that they watched them frequently. A group of 200 randomly selected college students was surveyed; 105 stated that they never watched the home shopping channels, 72 stated that they watched them rarely, 13 stated that they watched them occasionally, and 10 stated that they watched them frequently. At $\alpha = 0.05$, can it be concluded that the college students differ in their preference for the home shopping channels?

Source: Based on information obtained from *USA TODAY* Snapshots.

14. **Ways to Get to Work** The 2010 Census indicated the following percentages for means of commuting to work for workers over 15 years of age.

Alone	76.6
Carpooling	9.7
Public	4.9
Walked	2.8
Other	1.7
Worked at home	4.3

A random sample of workers found that 320 drove alone, 100 carpooled, 30 used public transportation, 20 walked, 10 used other forms of transportation, and 20 worked at home. Is there sufficient evidence to conclude that the proportions of workers using each type of transportation differ from those in the Census report? Use $\alpha = 0.05$.

Source: U.S. Census Bureau, *Washington Observer-Reporter.*

15. **Favorite Ice Cream Flavor** A survey of randomly selected women and randomly selected men asked what their favorite ice cream flavor was. The results are shown. At $\alpha = 0.05$, can it be concluded that the favorite flavor is independent of gender?

	Flavor			
	Vanilla	Chocolate	Strawberry	Other
Women	62	36	10	2
Men	49	37	5	9

16. **Types of Pizzas Purchased** A pizza shop owner wishes to determine if the type of pizza a person selects is related to the age of the individual. The data obtained from a sample are shown. At $\alpha = 0.10$, is the age of the purchaser related to the type of pizza ordered? Use the *P*-value method.

	Type of pizza			
Age	**Plain**	**Pepperoni**	**Mushroom**	**Double cheese**
10–19	12	21	39	71
20–29	18	76	52	87
30–39	24	50	40	47
40–49	52	30	12	28

17. **Pennant Colors Purchased** A survey at a ballpark shows the following selection of pennants sold to randomly selected fans. The data are presented here. At $\alpha = 0.10$, is the color of the pennant purchased independent of the gender of the individual?

	Blue	**Yellow**	**Red**
Men	519	659	876
Women	487	702	787

18. **Tax Credit Refunds** In a survey of randomly selected children ages 8 through 11 years, data were obtained as to what they think their parents should do with the money from a $400 tax credit.

	Keep it for themselves	**Give it to their children**	**Don't know**
Girls	162	132	6
Boys	147	147	6

At $\alpha = 0.10$, is there a relationship between the feelings of the children and the gender of the children?

Source: Based on information from *USA TODAY* Snapshot.

19. **Employment Satisfaction** A survey of 60 randomly selected men and 60 randomly selected women asked if they would be happy spending the rest of their careers with their present employers. The results are shown. At $\alpha = 0.10$, can it be concluded that the proportions are equal? If they are not equal, give a possible reason for the difference.

	Yes	**No**	**Undecided**
Men	40	15	5
Women	36	9	15

Source: Based on information from a Maritz Poll.

Critical Thinking Challenges

1. **Random Digits** Use your calculator or the MINITAB random number generator to generate 100 two-digit random numbers. Make a grouped frequency distribution, using the chi-square goodness-of-fit test to see if the distribution is random. To do this, use an expected frequency of 10 for each class. Can it be concluded that the distribution is random? Explain.
2. **Lottery Numbers** Simulate the state lottery by using your calculator or MINITAB to generate 100 three-digit random numbers. Group these numbers 000–099, 100–199, etc. Use the chi-square goodness-of-fit test to see if the numbers are random. The expected frequency for each class should be 10. Explain why.
3. **M&M's Colors** Purchase a bag of M&M's candy and count the number of pieces of each color. Using the information as your sample, state a hypothesis for the distribution of colors, and compare your hypothesis to H_0: The distribution of colors of M&M's candy is 13% brown, 13% red, 14% yellow, 16% green, 20% orange, and 24% blue.

Data Projects

Use a significance level of 0.05 for all tests below.

1. **Business and Finance** Many of the companies that produce multicolored candy will include on their website information about the production percentages for the various colors. Select a favorite multicolored candy. Find out what percentage of each color is produced. Open up a bag of the candy, noting how many of each color are in the bag (be careful to count them before you eat them). Is the bag distributed as expected based on the production percentages? If no production percentages can be found, test to see if the colors are uniformly distributed.
2. **Sports and Leisure** Use a local (or favorite) basketball, football, baseball, or hockey team as the data set. For the most recently completed season, note the team's home record for wins and losses. Test to see whether home field advantage is independent of sport.
3. **Technology** Use the data collected in data project 3 of Chapter 2 regarding song genres. Do the data indicate that songs are uniformly distributed among the genres?

4. **Health and Wellness** Research the percentages of each blood type that the Red Cross states are in the population. Now use your class as a sample. For each student note the blood type. Is the distribution of blood types in your class as expected based on the Red Cross percentages?
5. **Politics and Economics** Research the distribution (by percent) of registered Republicans, Democrats, and Independents in your state. Use your class as a sample. For each student, note the party affiliation. Is the distribution as expected based on the percentages for your state? What might be problematic about using your class as a sample for this exercise?
6. **Your Class** Conduct a classroom poll to determine which of the following sports each student likes best: baseball, football, basketball, hockey, or NASCAR. Also, note the gender of the individual. Is preference for sport independent of gender?

Answers to Applying the Concepts

Section 11–1 Skittles Color Distribution

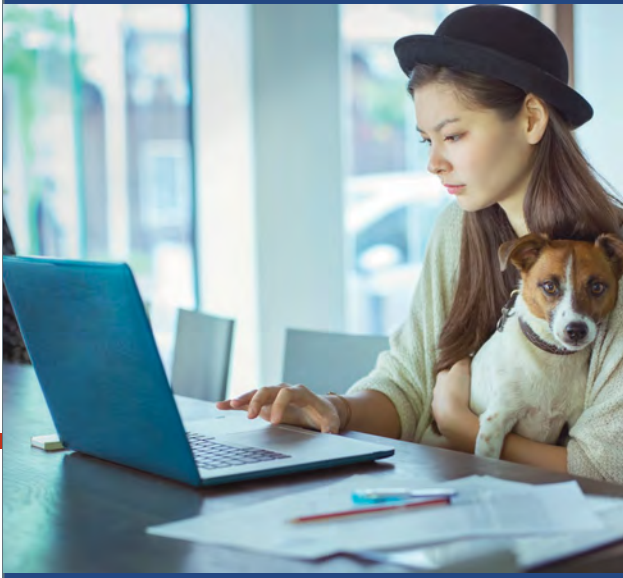
1. The variables are qualitative, and we have the counts for each category.
2. We can use a chi-square goodness-of-fit test.
3. There are a total of 233 candies, so we would expect 46.6 of each color. Our test statistic is $\chi^2 = 1.442$.
4. H_0: The colors are equally distributed.
   H_1: The colors are not equally distributed.
5. There are $5 - 1 = 4$ degrees of freedom for the test. The critical value depends on the choice of significance level. At the 0.05 significance level, the critical value is 9.488.
6. Since $1.442 < 9.488$, we fail to reject the null hypothesis. There is not enough evidence to conclude that the colors are not equally distributed.

Section 11–2 Satellite Dishes in Restricted Areas

1. We compare the P-value to the significance level of 0.05 to check if the null hypothesis should be rejected.
2. The significance level gives the probability of a type I error.
3. This is a right-tailed test, since chi-square tests of independence are always right-tailed.
4. You cannot tell how many rows and columns there were just by looking at the degrees of freedom.
5. Increasing the sample size does not increase the degrees of freedom, since the degrees of freedom are based on the number of rows and columns.
6. We will reject the null hypothesis. There are a number of cells where the observed and expected frequencies are quite different.
7. If the significance level were initially set at 0.10, we would still reject the null hypothesis.
8. No, the chi-square value does not tell us which cells have observed and expected frequencies that are very different.

12

Analysis of Variance



© Caiaimage/Glow Images RF

STATISTICS TODAY

Can Bringing Your Dog to Work Reduce Stress?

A study done by Virginia Commonwealth University researchers found that employees who brought their dogs to work had reduced job-related stress. The researchers selected a business that allowed employees to bring their dogs to work. Then they divided the employees into three groups: those who brought their dogs to work, those who owned dogs but did not bring them to work, and those who did not own any pets. For one week, the researchers measured the levels of the stress hormone cortisol in the workers' saliva and used surveys to gauge the stress levels of the three groups of employees.

After the week was up, the researchers compared the average levels of cortisol and the stress survey results of the three groups. Since there were three groups, the statistical methods that have been presented so far cannot be used to compare means of three groups. This chapter will show how to compare three or more means. See Statistics Today—Revisited at the end of the chapter to see how this might be done.

OUTLINE

Introduction
12–1 One-Way Analysis of Variance
12–2 The Scheffé Test and the Tukey Test
12–3 Two-Way Analysis of Variance
Summary

OBJECTIVES

After completing this chapter, you should be able to

1. Use the one-way ANOVA technique to determine if there is a significant difference among three or more means.
2. Determine which means differ, using the Scheffé or Tukey test if the null hypothesis is rejected in the ANOVA.
3. Use the two-way ANOVA technique to determine if there is a significant difference in the main effects or interaction.

Historical Note

The methods of analysis of variance were developed by R. A. Fisher in the early 1920s.

Introduction

The F test, used to compare two variances as shown in Chapter 9, can also be used to compare three or more means. This technique is called *analysis of variance,* or *ANOVA.* It is used to test claims involving three or more means. (*Note:* The F test can also be used to test the equality of two means. But since it is equivalent to the t test in this case, the t test is usually used instead of the F test when there are only two means.) For example, suppose a researcher wishes to see whether the means of the time it takes three groups of students to solve a computer problem using HTML, Java, and PHP are different. The researcher will use the ANOVA technique for this test. The z and t tests should not be used when three or more means are compared, for reasons given later in this chapter.

For three groups, the F test can show only whether a difference exists among the three means. It cannot reveal where the difference lies—that is, between $\overline{X}_1$ and $\overline{X}_2$, or $\overline{X}_1$ and $\overline{X}_3$, or $\overline{X}_2$ and $\overline{X}_3$. If the F test indicates that there is a difference among the means, other statistical tests are used to find where the difference exists. The most commonly used tests are the Scheffé test and the Tukey test, which are also explained in this chapter.

The analysis of variance that is used to compare three or more means is called a *one-way analysis of variance* since it contains only one variable. In the previous example, the variable is the type of computer language used. The analysis of variance can be extended to studies involving two variables, such as type of computer language used and mathematical background of the students. These studies involve a *two-way analysis of variance.* Section 12–3 explains the two-way analysis of variance.

12–1 One-Way Analysis of Variance

OBJECTIVE 1

Use the one-way ANOVA technique to determine if there is a significant difference among three or more means.

When an F test is used to test a hypothesis concerning the means of three or more populations, the technique is called **analysis of variance** (commonly abbreviated as **ANOVA**).

The **one-way analysis of variance** test is used to test the equality of three or more means using sample variances.

The procedure used in this section is called the one-way analysis of variance because there is only one independent variable that distinguishes between the different populations in the study. The independent variable is also called a *factor*.

At first glance, you might think that to compare the means of three or more samples, you can use multiple t tests comparing two means at a time. But there are several reasons why the t test should not be done.

First, when you are comparing two means at a time, the rest of the means under study are ignored. With the F test, all the means are compared simultaneously. Second, when you are comparing two means at a time and making all pairwise comparisons, the probability of rejecting the null hypothesis when it is true is increased, since the more t tests that are conducted, the greater is the likelihood of getting significant differences by chance alone. Third, the more means there are to compare, the more t tests are needed. For example, for the comparison of 3 means two at a time, 3 t tests are required. For the comparison of 5 means two at a time, 10 tests are required. And for the comparison of 10 means two at a time, 45 tests are required.

As the number of populations to be compared increases, the probability of making a type I error using multiple t tests for a given level of significance α also increases. To address this problem, the technique of analysis of variance is used. This technique involves a comparison of two estimates of the same population variance.

Recall that the characteristics of the F distribution are as follows:

1. The values of F cannot be negative, because variances are always positive or zero.
2. The distribution is positively skewed.
3. The mean value of F is approximately equal to 1.
4. The F distribution is a family of curves based on the degrees of freedom of the variance of the numerator and the degrees of freedom of the variance of the denominator.

Even though you are comparing three or more means in this use of the F test, *variances* are used in the test instead of means.

With the F test, two different estimates of the population variance are made. The first estimate is called the **between-group variance,** and it involves finding the variance of the means. The second estimate, the **within-group variance,** is made by computing the variance using all the data and is not affected by differences in the means. If there is no difference in the means, the between-group variance estimate will be approximately equal to the within-group variance estimate, and the F test value will be approximately equal to one. The null hypothesis will not be rejected. However, when the means differ significantly, the between-group variance will be much larger than the within-group variance; the F test value will be significantly greater than one; and the null hypothesis will be rejected. Since variances are compared, this procedure is called *analysis of variance* (ANOVA).

The formula for the F test is

$$F = \frac{\text{variance between groups}}{\text{variance within groups}}$$

The variance between groups measures the differences in the means that result from the different treatments given to each group. To calculate this value, it is necessary to find the *grand mean* $\overline{X}_{\text{GM}}$, which is the mean of all the values in all of the samples. The formula for the grand mean is

$$\overline{X}_{\text{GM}} = \frac{\Sigma X}{N}$$

This value is used to find the between-group variance s_B^2. This is the variance among the means using the sample sizes as weights.

The formula for the between-group variance, denoted by s_B^2, is

$$s_B^2 = \frac{\Sigma n_i(\overline{X}_i - \overline{X}_{\text{GM}})^2}{k-1}$$

where k = number of groups
n_i = sample size
$\overline{X}_i$ = sample mean

This formula can be written out as

$$s_B^2 = \frac{n_1(\overline{X}_1 - \overline{X}_{\text{GM}})^2 + n_2(\overline{X}_2 - \overline{X}_{\text{GM}})^2 + \cdots + n_k(\overline{X}_k - \overline{X}_{\text{GM}})^2}{k-1}$$

Next find the within group variance, denoted by s_W^2. The formula finds the overall variance by calculating a weighted average of the individual variances. It does not involve using differences of means. The formula for the within-group variance is

$$s_W^2 = \frac{\Sigma(n_i - 1)s_i^2}{\Sigma(n_i - 1)}$$

where n_i = sample size
s_i^2 = variance of sample

This formula can be written out as

$$s_W^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \cdots + (n_k - 1)s_k^2}{(n_1 - 1) + (n_2 - 1) + \cdots + (n_k - 1)}$$

Finally, the F test value is computed. The formula can now be written using the symbols s_B^2 and s_W^2.

The formula for the F test for one-way analysis of variance is

$$F = \frac{s_B^2}{s_W^2}$$

where s_B^2 = between-group variance
s_W^2 = within-group variance

For a test of the difference among three or more means, the following hypotheses can be used.

$$H_0: \mu_1 = \mu_2 = \cdots = \mu_n$$
$$H_1: \text{At least one mean is different from the others}$$

The alternative hypothesis means that at least one mean differs from one or more of the other means and as previously stated, a significant test value means that there is a high probability that this difference in means is not due to chance, but it does not indicate where the difference lies.

The degrees of freedom for this F test are d.f.N. $= k - 1$, where k is the number of groups, and d.f.D. $= N - k$, where N is the sum of the sample sizes of the groups $N = n_1 + n_2 + \cdots + n_k$. The sample sizes need not be equal. The F test to compare means is always right-tailed.

The results of the one-way analysis of variance can be summarized by placing them in an **ANOVA summary table.** The numerator of the fraction of the s_B^2 term is called the **sum of squares between groups,** denoted by SS_B. The numerator of the s_W^2 term is called the **sum of squares within groups,** denoted by SS_W. This statistic is also called the *sum of squares for the error.* SS_B is divided by d.f.N. to obtain the between-group variance. SS_W is divided by $N - k$ to obtain the within-group or error variance. These two variances are sometimes called **mean squares,** denoted by MS_B and MS_W. These terms are used to summarize the analysis of variance and are placed in a summary table, as shown in Table 12–1.

TABLE 12–1 Analysis of Variance Summary Table

Source	Sum of squares	d.f.	Mean square	F
Between	SS_B	$k - 1$	MS_B	
Within (error)	SS_W	$N - k$	MS_W	
Total				

Unusual Stat

The *Journal of the American College of Nutrition* reports that a study found no correlation between body weight and the percentage of calories eaten after 5:00 P.M.

In the table,

$$\text{SS}_B = \text{sum of squares between groups}$$
$$\text{SS}_W = \text{sum of squares within groups}$$
$$k = \text{number of groups}$$
$$N = n_1 + n_2 + \cdots + n_k = \text{sum of sample sizes for groups}$$
$$\text{MS}_B = \frac{\text{SS}_B}{k - 1}$$
$$\text{MS}_W = \frac{\text{SS}_W}{N - k}$$
$$F = \frac{\text{MS}_B}{\text{MS}_W}$$

To use the F test to compare two or more means, the following assumptions must be met.

Assumptions for the *F* Test for Comparing Three or More Means

1. The populations from which the samples were obtained must be normally or approximately normally distributed.
2. The samples must be independent of one another.
3. The variances of the populations must be equal.
4. The samples must be simple random samples, one from each of the populations.

In this book, the assumptions will be stated in the exercises; however, when encountering statistics in other situations, you must check to see that these assumptions have been met before proceeding.

The steps for computing the F test value for the ANOVA are summarized in this Procedure Table.

Procedure Table

Finding the *F* Test Value for the Analysis of Variance

Step 1 Find the mean and variance of each sample.

$$(\overline{X}_1, s_1^2), (\overline{X}_2, s_2^2), \ldots, (\overline{X}_k, s_k^2)$$

Step 2 Find the grand mean.

$$\overline{X}_{GM} = \frac{\Sigma X}{N}$$

Step 3 Find the between-group variance.

$$s_B^2 = \frac{\Sigma n_i(\overline{X}_i - \overline{X}_{GM})^2}{k - 1}$$

Step 4 Find the within-group variance.

$$s_W^2 = \frac{\Sigma(n_i - 1)s_i^2}{\Sigma(n_i - 1)}$$

Step 5 Find the *F* test value.

$$F = \frac{s_B^2}{s_W^2}$$

The degrees of freedom are

$$\text{d.f.N.} = k - 1$$

where k is the number of groups, and

$$\text{d.f.D.} = N - k$$

where N is the sum of the sample sizes of the groups

$$N = n_1 + n_2 + \cdots + n_k$$

The one-way analysis of variance follows the regular five-step hypothesis-testing procedure.

Step 1 State the hypotheses.

Step 2 Find the critical values.

Step 3 Compute the test value.

Step 4 Make the decision.

Step 5 Summarize the results.

Examples 12–1 and 12–2 illustrate the computational procedure for the ANOVA technique for comparing three or more means, and the steps are summarized in the Procedure Table.

EXAMPLE 12–1 Miles per Gallon

A researcher wishes to see if there is a difference in the fuel economy for city driving for three different types of automobiles: small automobiles, sedans, and luxury automobiles. He randomly samples four small automobiles, five sedans, and three luxury automobiles. The miles per gallon for each is shown. At $\alpha = 0.05$, test the claim that there is no difference among the means. The data are shown.

Small	Sedans	Luxury
36	43	29
44	35	25
34	30	24
35	29	
	40	

Source: U.S. Environmental Protection Agency.

Step 1 State the hypotheses and identify the claim.

H_0: $\mu_1 = \mu_2 = \mu_3$ (claim)

H_1: At least one mean is different from the others

Step 2 Find the critical value.

$$N = 12 \qquad k = 3$$
$$\text{d.f.N.} = k - 1 = 3 - 1 = 2$$
$$\text{d.f.D.} = N - k = 12 - 3 = 9$$

The critical value from Table H in Appendix A with $\alpha = 0.05$ is 4.26.

Step 3 Compute the test value.

a. Find the mean and variance for each sample. (Use the formulas in Chapter 3.)

For the small cars:	$\overline{X} = 37.25$	$s^2 = 20.917$
For the sedans:	$\overline{X} = 35.4$	$s^2 = 37.3$
For the luxury cars:	$\overline{X} = 26$	$s^2 = 7$

b. Find the grand mean.

$$\overline{X}_{\text{GM}} = \frac{\Sigma X}{N} = \frac{36 + 44 + 34 + \cdots + 24}{12} = \frac{404}{12} = 33.667$$

c. Find the between-group variance.

$$s_B^2 = \frac{\Sigma n(\overline{X}_i - \overline{X}_{\text{GM}})^2}{k - 1}$$

$$= \frac{4(37.25 - 33.667)^2 + 5(35.4 - 33.667)^2 + 3(26 - 33.667)^2}{3 - 1}$$

$$= \frac{242.717}{2} = 121.359$$

d. Find the within-group variance.

$$s_W^2 = \frac{\Sigma(n_i - 1)s_i^2}{\Sigma(n_i - 1)} = \frac{(4 - 1)(20.917) + (5 - 1)(37.3) + (3 - 1)7}{(4 - 1) + (5 - 1) + (3 - 1)}$$

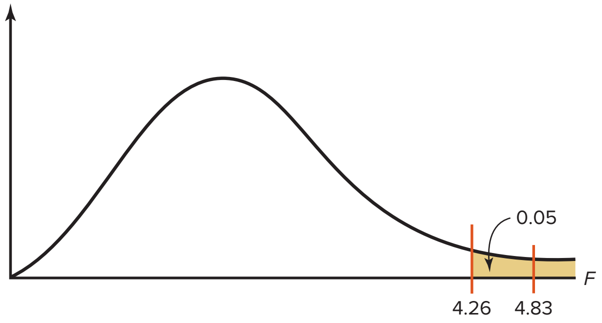
$$= \frac{225.951}{9} = 25.106$$

e. Find the F test value.

$$F = \frac{s_B^2}{s_W^2} = \frac{121.359}{25.106} = 4.83$$

Step 4 Make the decision. The test value 4.83 > 4.26, so the decision is to reject the null hypothesis. See Figure 12–1.

FIGURE 12–1 Critical Value and Test Value for Example 12–1



Step 5 Summarize the results. There is enough evidence to conclude that at least one mean is different from the others.

The ANOVA summary table is shown in Table 12–2.

TABLE 12–2 Analysis of Variance Summary Table for Example 12–1

Source	Sum of squares	d.f.	Mean square	F
Between	242.717	2	121.359	4.83
Within (error)	225.954	9	25.106	
Total	468.671	11		

The P-values for ANOVA are found by using the same procedure shown in Section 9–5. For Example 12–1, the F test value is 4.83. In Table H with d.f.N. = 2 and d.f.D. = 9, the F test value falls between $\alpha = 0.025$ with an F value of 5.71 and $\alpha = 0.05$ with an F value of 4.26. Hence, $0.025 < P\text{-value} < 0.05$. In this case, the null hypothesis is rejected at $\alpha = 0.05$ since the P-value < 0.05. The TI-84 P-value is 0.0375.

EXAMPLE 12–2 Tall Buildings

A researcher wishes to see if there is a difference in the number of stories in the tall buildings of Chicago, Houston, and New York City. The researcher randomly selects five buildings in each city and records the number of stories in each building. The data are shown. At $\alpha = 0.05$, can it be concluded that there is a significant difference in the mean number of stories in the tall buildings in each city?

Chicago	Houston	New York City
98	53	85
54	52	67
60	45	75
57	41	52
83	36	94
49	34	42

Source: The World Almanac and Book of Facts

SOLUTION

Step 1 State the hypotheses and identify the claim.

H_0: $\mu_1 = \mu_2 = \mu_3$

H_1: At least one mean is different from the others (claim).

Step 2 Find the critical value. Since $k = 3$, $N = 18$, and $\alpha = 0.05$,

d.f.N. $= k - 1 = 3 - 1 = 2$

d.f.D. $= N - k = 18 - 3 = 15$

The critical value is 3.68.

Step 3 Compute the test value.

a. Find the mean and variance of each sample. The mean and variance for each sample are

Chicago $\overline{X}_1 = 66.8$ $s_1^2 = 371.8$

Houston $\overline{X}_2 = 43.5$ $s_2^2 = 63.5$

New York $\overline{X}_3 = 69.2$ $s_3^2 = 387.7$

b. Find the grand mean.

$$\overline{X}_{\text{GM}} = \frac{\Sigma X}{N} = \frac{98 + 54 + 60 + \cdots + 42}{18} = \frac{1077}{18} = 59.8$$

c. Find the between-group variance.

$$s_B^2 = \frac{\Sigma n_i(\overline{X}_i - \overline{X}_{\text{GM}})^2}{k-1}$$

$$= \frac{6(66.8-59.8)^2 + 6(43.5-59.8)^2 + 6(69.2-59.8)^2}{3-1}$$

$$= \frac{2418.3}{2} = 1209.15$$

d. Find the within-group variance.

$$s_W^2 = \frac{\Sigma(n_i-1)s_i^2}{\Sigma(n_i-1)}$$

$$= \frac{(6-1)(371.8) + (6-1)(63.5) + (6-1)(387.7)}{(6-1)+(6-1)+(6-1)} = \frac{4115}{15}$$

$$= 274.33$$

e. Find the *F* test value.

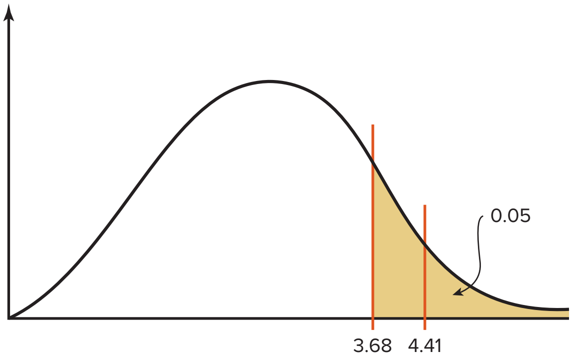
$$F = \frac{s_B^2}{s_W^2} = \frac{1209.15}{274.33} = 4.41$$

Interesting Facts

The weight of 1 cubic foot of wet snow is about 10 pounds while the weight of 1 cubic foot of dry snow is about 3 pounds.

Step 4 Make the decision. Since 4.41 > 3.68, the decision is to reject the null hypothesis. See Figure 12–2.

FIGURE 12–2 Critical Value and Test Value for Example 12–2



Step 5 Summarize the results. There is enough evidence to support the claim that at least one mean is different from the others. The ANOVA summary table for this example is shown in Table 12–3.

TABLE 12–3 Analysis of Variance Summary Table for Example 12–2

Source	Sum of squares	d.f.	Mean square	*F*
Between	2418.3	2	1209.15	4.41
Within	4115	15		
Total	6533.3	17		

In this case, 4.41 falls between 4.77 and 3.68, which corresponds to 0.025 at the 0.05 level. Hence, 0.025 < *P*-value < 0.05. Since the *P*-value is less than 0.05, the decision is to reject the null hypothesis. The *P*-value obtained from the calculator is 0.031.

When the null hypothesis is rejected in ANOVA, it only means that at least one mean is different from the others. To locate the difference or differences among the means, it is necessary to use other tests such as the Tukey or the Scheffé test.

Applying the Concepts 12–1

Colors That Make You Smarter

The following set of data values was obtained from a study of people's perceptions on whether the color of a person's clothing is related to how intelligent the person looks. The subjects rated the person's intelligence on a scale of 1 to 10. Randomly selected group 1 subjects were shown people with clothing in shades of blue and gray. Randomly selected group 2 subjects were shown people with clothing in shades of brown and yellow. Randomly selected group 3 subjects were shown people with clothing in shades of pink and orange. The results follow.

Group 1	Group 2	Group 3
8	7	4
7	8	9
7	7	6
7	7	7
8	5	9
8	8	8
6	5	5
8	8	8
8	7	7
7	6	5
7	6	4
8	6	5
8	6	4

1. Use ANOVA to test for any significant differences between the means.
2. What is the purpose of this study?
3. Explain why separate *t* tests are not accepted in this situation.

See page 683 for the answers.

Exercises 12–1

1. What test is used to compare three or more means?
2. State three reasons why multiple *t* tests cannot be used to compare three or more means.
3. What are the assumptions for ANOVA?
4. Define between-group variance and within-group variance.
5. State the hypotheses used in the ANOVA test.
6. When there is no significant difference among three or more means, the value of *F* will be close to what number?

For Exercises 7 through 20, assume that all variables are normally distributed, that the samples are independent, that the population variances are equal, and that the samples are simple random samples, one from each of the populations. Also, for each exercise, perform the following steps.

a. State the hypotheses and identify the claim.
b. Find the critical value.
c. Compute the test value.
d. Make the decision.
e. Summarize the results, and explain where the differences in the means are.

Use the traditional method of hypothesis testing unless otherwise specified.

7. **Leading Businesses** The following data show the yearly budgets for leading business sectors in the United States. At $\alpha = 0.05$, is there a significant difference in the mean budgets of the business sectors? The data are in thousands of dollars.

Beverages	Electronics	Food producers	Supportive services
170	46	59	56
128	24	58	37
19	18	33	19
16	14	31	19
12	13	28	17
11	12	22	15
10	10	16	15

Source: Financial Times 500.

8. **Sodium Contents of Foods** The amount of sodium (in milligrams) in one serving for a random sample of three different kinds of foods is listed. At the 0.05 level of significance, is there sufficient evidence to conclude that a difference in mean sodium amounts exists among condiments, cereals, and desserts?

Condiments	Cereals	Desserts
270	260	100
130	220	180
230	290	250
180	290	250
80	200	300
70	320	360
200	140	300
		160

Source: The Doctor's Pocket Calorie, Fat, and Carbohydrate Counter.

9. **Hybrid Vehicles** A study was done before the recent surge in gasoline prices to compare the cost to drive 25 miles for different types of hybrid vehicles. The cost of a gallon of gas at the time of the study was approximately $2.50. Based on the information given for different models of hybrid cars, trucks, and SUVs, is there sufficient evidence to conclude a difference in the mean cost to drive 25 miles? Use $\alpha = 0.05$. (The information in this exercise will be used in Exercise 3 in Section 12–2.)

Hybrid cars	Hybrid SUVs	Hybrid trucks
2.10	2.10	3.62
2.70	2.42	3.43
1.67	2.25	
1.67	2.10	
1.30	2.25	

Source: www.fueleconomy.com

10. **Healthy Eating** Americans appear to be eating healthier. Between 1970 and 2013 the per capita consumption of broccoli increased 1200.5% from 0.5 to 6.4 pounds. A nutritionist followed a group of people randomly assigned to one of three groups and noted their monthly broccoli intake (in pounds). At $\alpha = 0.05$, is there a difference in means?

Group A	Group B	Group C
2.0	2.0	3.7
1.5	1.5	2.5
0.75	4.0	4.0
1.0	3.0	5.1
1.3	2.5	3.8
3.0	2.0	2.9

Source: World Almanac.

11. **Movie Theater Attendance** The data shown are the weekly admissions, in millions, of people attending movie theaters over three different time periods. At $\alpha = 0.05$, is there a difference in the means for the weekly attendance for these time periods?

1950–1974	1975–1990	1991–2000
58.0	17.1	23.3
39.9	19.9	26.6
25.1	19.6	27.7
19.8	20.3	26.5
17.7	22.9	25.8

Source: Motion Picture Association.

12. **Weight Gain of Athletes** A researcher wishes to see whether there is any difference in the weight gains of athletes following one of three special diets. Athletes are randomly assigned to three groups and placed on the diet for 6 weeks. The weight gains (in pounds) are shown here. At $\alpha = 0.05$, can the researcher conclude that there is a difference in the diets?

Diet A	Diet B	Diet C
3	10	8
6	12	3
7	11	2
4	14	5
	8	
	6	

A computer printout for this problem is shown. Use the P-value method and the information in this printout to test the claim. (The information in this exercise will be used in Exercise 4 of Section 12–2.)

Computer Printout for Exercise 12

```
ANALYSIS OF VARIANCE SOURCE TABLE
Source       df  Sum of Squares  Mean Square      F    P-value
Bet Groups    2         101.095       50.548  7.740    0.00797
W/I Groups   11          71.833        6.530
Total        13         172.929
DESCRIPTIVE STATISTICS
Condit        N           Means       St Dev
diet A        4           5.000        1.826
diet B        6          10.167        2.858
diet C        4           4.500        2.646
```

13. **Expenditures per Pupil** The per-pupil costs (in thousands of dollars) for cyber charter school tuition for school districts in three areas of southwestern Pennsylvania are shown. At $\alpha = 0.05$, is there a difference in the means? If so, give a possible reason for the difference. (The information in this exercise will be used in Exercise 5 of Section 12–2.)

Area I	Area II	Area III
6.2	7.5	5.8
9.3	8.2	6.4
6.8	8.5	5.6
6.1	8.2	7.1
6.7	7.0	3.0
6.9	9.3	3.5

Source: Tribune-Review.

14. **Cell Phone Bills** The average local cell phone monthly bill is $50.07. A random sample of monthly bills from three different providers is listed below. At $\alpha = 0.05$,

is there a difference in mean bill amounts among providers?

Provider X	Provider Y	Provider Z
48.20	105.02	59.27
60.59	85.73	65.25
72.50	61.95	70.27
55.62	75.69	42.19
89.47	82.11	52.34

Source: World Almanac.

15. **Air Quality** The data show the particulate matter in micrograms per cubic meter for a sample of large cities on three continents. At $\alpha = 0.10$, is there a difference in the mean particulate matter among these continents?

Asia	Europe	North America
83	24	21
46	36	27
118	34	54
154	27	17
50	22	16
127	42	
	29	
	26	
	30	

Source: The World Almanac and Book of Facts.

16. **Annual Child Care Costs** Annual child care costs for infants are considerably higher than for older children. At $\alpha = 0.05$, can you conclude a difference in mean infant day care costs for different regions of the United States? (Annual costs per infant are given in dollars.)

New England	Midwest	Southwest
10,390	9,449	7,644
7,592	6,985	9,691
8,755	6,677	5,996
9,464	5,400	5,386
7,328	8,372	

Source: www.naccrra.org (National Association of Child Care Resources and Referral Agencies: "Breaking the Piggy Bank").

17. **Microwave Oven Prices** A research organization tested microwave ovens. At $\alpha = 0.10$, is there a significant difference in the average prices of the three types of oven?

	Watts	
1000	**900**	**800**
270	240	180
245	135	155
190	160	200
215	230	120
250	250	140
230	200	180
	200	140
	210	130

A computer printout for this exercise is shown. Use the *P*-value method and the information in this printout to test the claim. (The information in this exercise will be used in Exercise 6 of Section 12–2.)

Computer Printout for Exercise 17

```
ANALYSIS OF VARIANCE SOURCE TABLE
Source      df  Sum of Squares  Mean Square       F   P-value
Bet Groups   2       21729.735    10864.867  10.118   0.00102
W/I Groups  19       20402.083     1073.794
Total       21       42131.818
DESCRIPTIVE STATISTICS
Condit       N           Means       St Dev
 1000        6         233.333        28.23
  900        8         203.125        39.36
  800        8         155.625        28.21
```

18. **Calories in Fast-Food Sandwiches** Three popular fast-food restaurant franchises specializing in burgers were surveyed to find out the number of calories in their frequently ordered sandwiches. At the 0.05 level of significance, can it be concluded that a difference in mean number of calories per burger exists? (The information in this exercise will be used for Exercise 7 in Section 12–2.)

FF#1	FF#2	FF#3
970	1010	740
880	970	540
840	920	510
710	850	510
	820	

Source: www.fatcalories.com

19. **Number of Pupils in a Class** A large school district has several middle schools. Three schools were randomly chosen, and four classes were selected from each. The numbers of pupils in each class are shown here. At $\alpha = 0.10$, is there sufficient evidence that the mean number of students per class differs among schools?

MS 1	MS 2	MS 3
21	28	25
25	22	20
19	25	23
17	30	22

20. **Average Debt of College Graduates** Kiplinger's listed the top 100 public colleges based on many factors. From that list, here is the average debt at graduation for various schools in four selected states. At $\alpha = 0.05$, can it be concluded that the average debt at graduation differs for these four states?

New York	Virginia	California	Pennsylvania
14,734	14,524	13,171	18,105
16,000	15,176	14,431	17,051
14,347	12,665	14,689	16,103
14,392	12,591	13,788	22,400
12,500	18,385	15,297	17,976

Source: www.Kiplinger.com

Technology Step by Step

TI-84 Plus
Step by Step

One-Way Analysis of Variance (ANOVA)

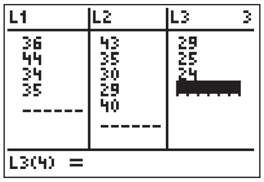
1. Enter the data into L_1, L_2, L_3, etc.
2. Press **STAT** and move the cursor to TESTS.
3. Press **H (ALPHA ^)** for ANOVA(.
4. Type each list followed by a comma. End with) and press **ENTER.**

Example TI12–1

Test the claim H_0: $\mu_1 = \mu_2 = \mu_3 =$ at $\alpha = 0.05$ for these data from Example 12–1.

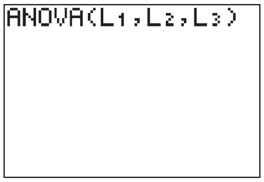
Small	Sedans	Luxury
36	43	29
44	35	25
34	30	24
35	29	
	40	

Input



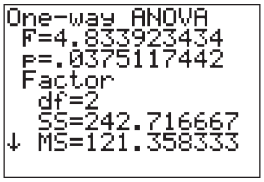
```
L1      L2      L3     3
 36      43      29
 44      35      25
 34      30      24
 35      29      ------
 ------  40
         ------
L3(4) =
```

Input



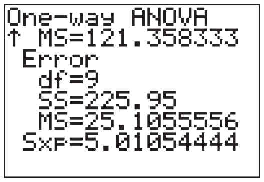
```
ANOVA(L1,L2,L3)
```

Output



```
One-way ANOVA
 F=4.833923434
 p=.0375117442
 Factor
  df=2
  SS=242.716667
↓ MS=121.358333
```

Output



```
One-way ANOVA
↑ MS=121.358333
 Error
  df=9
  SS=225.95
  MS=25.1055556
 Sxp=5.01054444
```

The F test value is 4.833923434. The P-value is 0.0375117442, which is significant at $\alpha = 0.05$. The factor variable has

d.f. = 2
SS = 242.716667
MS = 121.358333

The error has

d.f. = 9
SS = 225.95
MS = 25.1055556

EXCEL
Step by Step

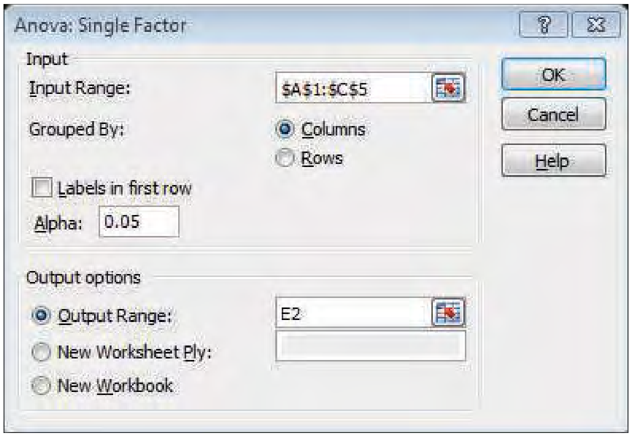
One-Way Analysis of Variance (ANOVA)

Example XL12–1

1. Enter the data below in columns A, B, and C of an Excel worksheet.

10	6	5
12	8	9
9	3	12
15	0	8
13	2	4

2. From the toolbar, select **Data,** then **Data Analysis.**
3. Select **Anova: Single Factor.**
4. Type in **A1:C5** in the **Input Range** box.
5. Check **Grouped By: Columns.**
6. Type **0.05** for the **Alpha level.**
7. Under the **Output options,** check the **Output Range** and type **E2**.
8. Click [OK].



Anova: Single Factor

Input

Input Range: $A$1:$C$5

Grouped By: ● Columns ○ Rows

☐ Labels in first row

Alpha: 0.05

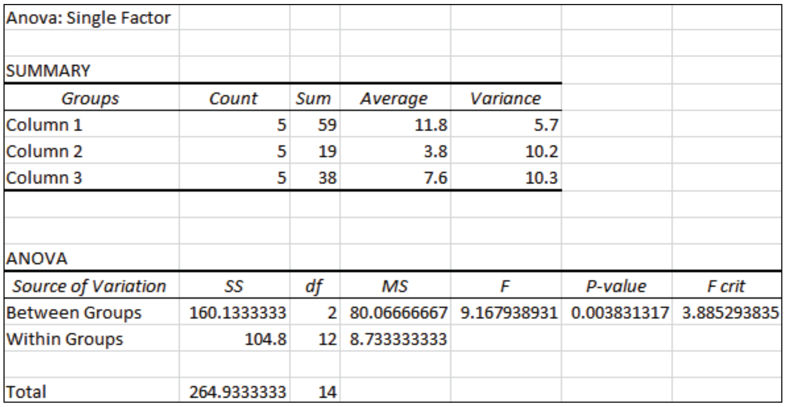
Output options

● Output Range: E2

○ New Worksheet Ply:

○ New Workbook

The results of the ANOVA are shown below.



Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
Column 1	5	59	11.8	5.7
Column 2	5	19	3.8	10.2
Column 3	5	38	7.6	10.3

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	160.1333333	2	80.06666667	9.167938931	0.003831317	3.885293835
Within Groups	104.8	12	8.733333333			
Total	264.9333333	14				

MINITAB

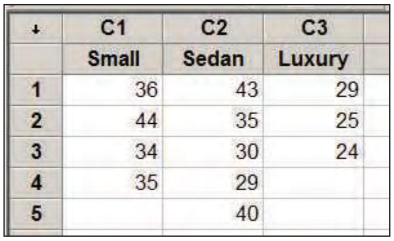
Step by Step

One-Way Analysis of Variance (ANOVA)

Example 12–1

Is there a difference in the average city MPG rating by type of vehicle?

1. Enter the MPG ratings in C1 Small, C2 Sedan, and C3 Luxury.



↓	C1	C2	C3
	Small	Sedan	Luxury
1	36	43	29
2	44	35	25
3	34	30	24
4	35	29	
5		40	

2. Select **Stat>ANOVA>One-way (unstacked).**
    - **a.** Select **Response data are in a separate column for each factor level** in the drop down menu.
    - **b.** Click in the Responses: dialog box.
    - **c.** Drag the mouse over the three columns of Observed counts.
    - **d.** Click on [Select].
    - **e.** Click [OK].

The results are displayed in the session window.

```
One-way ANOVA: Small, Sedan, Luxury
Source    DF       SS       MS       F       P
Factor     2    242.7    121.4    4.83   0.038
Error      9    226.0     25.1
Total     11    468.7

S = 5.011    R-Sq = 51.79%    R-Sq(adj) = 41.08%

                                    Individual 95% CIs for Mean Based on Pooled StDev
Level      N      Mean     StDev    -----------+------------+------------+------------+-
Small      4    37.250     4.573                                   (------------*------------)
Sedan      5    35.400     6.107                                (------------*------------)
Luxury     3    26.000     2.646    (-------------*-------------)
                                    ------------+------------+------------+------------+-
                                              24.0         30.0         36.0         42.0

Pooled StDev = 5.011
```

12–2 The Scheffé Test and the Tukey Test

OBJECTIVE 2

Determine which means differ, using the Scheffé or Tukey test if the null hypothesis is rejected in the ANOVA.

When the null hypothesis is rejected using the F test, the researcher may want to know where the difference among the means is. Several procedures have been developed to determine where the significant differences in the means lie after the ANOVA procedure has been performed. Among the most commonly used tests are the *Scheffé test* and the *Tukey test.*

Scheffé Test

To conduct the **Scheffé test,** you must compare the means two at a time, using all possible combinations of means. For example, if there are three means, the following comparisons must be done:

$$\overline{X}_1 \text{ versus } \overline{X}_2 \qquad \overline{X}_1 \text{ versus } \overline{X}_3 \qquad \overline{X}_2 \text{ versus } \overline{X}_3$$

Unusual Stat

According to the *British Medical Journal,* the body's circadian rhythms produce drowsiness during the midafternoon, matched only by the 2:00 A.M. to 7:00 A.M. period for sleep-related traffic accidents.

Formula for the Scheffé Test

$$F_s = \frac{(\overline{X}_i - \overline{X}_j)^2}{s_W^2[(1/n_i) + (1/n_j)]}$$

where $\overline{X}_i$ and $\overline{X}_j$ are the means of the samples being compared, n_i and n_j are the respective sample sizes, and s_W^2 is the within-group variance.

To find the critical value F' for the Scheffé test, multiply the critical value for the F test by $k - 1$:

$$F' = (k - 1)(\text{C.V.})$$

There is a significant difference between the two means being compared when the F test value, F_S, is greater than the critical value, F'. Example 12–3 illustrates the use of the Scheffé test.

EXAMPLE 12–3

Use the Scheffé test to test each pair of means in Example 12–1 to see if a significant difference exists between each pair of means. Use $\alpha = 0.05$.

SOLUTION

The F critical value for Example 12–1 is 4.26. Then the critical value for the individual tests with d.f.N. = 2 and d.f.D. = 9 is

$$F' = (k - 1)(\text{C.V.}) = (3 - 1)(4.26) = 8.52$$

a. For $\overline{X}_1$ versus $\overline{X}_2$,

$$F_S = \frac{(\overline{X}_1 - \overline{X}_2)^2}{s_W^2[(1/n_1) + (1/n_2)]} = \frac{(37.25 - 35.4)^2}{25.106\left(\frac{1}{4} + \frac{1}{5}\right)} = 0.30$$

Since $0.30 < 8.52$, the decision is that μ_1 is not significantly different from μ_2.

b. For $\overline{X}_1$ versus $\overline{X}_3$,

$$F_S = \frac{(\overline{X}_1 - \overline{X}_3)^2}{s_W^2[(1/n_1) + (1/n_3)]} = \frac{(37.25 - 26)^2}{25.106\left(\frac{1}{4} + \frac{1}{3}\right)} = 8.64$$

Since $8.64 > 8.52$, the decision is that μ_1 is significantly different from μ_3.

c. For $\overline{X}_2$ versus $\overline{X}_3$,

$$F_S = \frac{(\overline{X}_2 - \overline{X}_3)^2}{s_W^2[(1/n_2) + (1/n_3)]} = \frac{(35.4 - 26)^2}{25.106\left(\frac{1}{5} + \frac{1}{3}\right)} = 6.60$$

Since $6.60 < 8.64$, the decision is that μ_2 is not significantly different from μ_3. Hence, only the mean of the small cars is not equal to the mean of luxury cars.

On occasion, when the F test value is greater than the critical value, the Scheffé test may not show any significant differences in the pairs of means. This result occurs because the difference may actually lie in the average of two or more means when compared with the other mean. The Scheffé test can be used to make these types of comparisons, but the technique is beyond the scope of this book.

Tukey Test

The **Tukey test** can also be used after the analysis of variance has been completed to make pairwise comparisons between means when the groups have the same sample size. The symbol for the test value in the Tukey test is q.

Formula for the Tukey Test

$$q = \frac{\overline{X}_i - \overline{X}_j}{\sqrt{s_W^2/n}}$$

where $\overline{X}_i$ and $\overline{X}_j$ are the means of the samples being compared, n is the size of the samples, and s_W^2 is the within-group variance.

When the absolute value of q is greater than the critical value for the Tukey test, there is a significant difference between the two means being compared.

The critical value for the Tukey test is found using Table N in Appendix A, where k is the number of means in the original problem and v is the degrees of freedom for s_W^2, which is $N - k$. The value of k is found across the top row, and v is found in the left column.

EXAMPLE 12–4

Using the Tukey test, test each pair of means in Example 12–2 to see whether a specific difference exists, at $\alpha = 0.05$.

SOLUTION

a. For $\overline{X}_1$ versus $\overline{X}_2$,

$$q = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{s_W^2/n}} = \frac{66.8 - 43.5}{\sqrt{274.33/6}} = 3.446$$

b. For $\overline{X}_1$ versus $\overline{X}_3$,

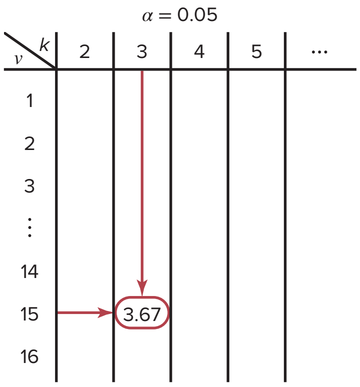
$$q = \frac{\overline{X}_1 - \overline{X}_3}{\sqrt{s_W^2/n}} = \frac{66.8 - 69.2}{\sqrt{274.33/6}} = -0.355$$

c. For $\overline{X}_2$ versus $\overline{X}_3$,

$$q = \frac{\overline{X}_2 - \overline{X}_3}{\sqrt{s_W^2/n}} = \frac{43.5 - 69.2}{\sqrt{274.33/6}} = -3.801$$

To find the critical value for the Tukey test, use Table N in Appendix A. The number of means k is found in the row at the top, and the degrees of freedom for s_W^2 are found in the left column (denoted by v). Since $k = 3$, d.f. $= 18 - 3 = 15$, and $\alpha = 0.05$, the critical value is 3.67. See Figure 12–3. Hence, the only q value that is greater in absolute value than the critical value is the one for the difference between $\overline{X}_2$ and $\overline{X}_3$. The conclusion, then, is that there is a significant difference in means for the Houston buildings and the New York City buildings.

FIGURE 12–3 Finding the Critical Value in Table N for the Tukey Test (Example 12–4)



You might wonder why there are two different tests that can be used after the ANOVA. Actually, there are several other tests that can be used in addition to the Scheffé and Tukey tests. It is up to the researcher to select the most appropriate test. The Scheffé test is the most general, and it can be used when the samples are of different sizes. Furthermore, the Scheffé test can be used to make comparisons such as the average of $\overline{X}_1$ and $\overline{X}_2$ compared with $\overline{X}_3$. However, the Tukey test is more powerful than the Scheffé test for making pairwise comparisons for the means. A rule of thumb for pairwise comparisons is to use the Tukey test when the samples are equal in size and the Scheffé test when the samples differ in size. This rule will be followed in this textbook.

Applying the Concepts 12–2

Colors That Make You Smarter

The following set of data values was obtained from a study of people's perceptions on whether the color of a person's clothing is related to how intelligent the person looks. The subjects rated the person's intelligence on a scale of 1 to 10. Randomly selected group 1 subjects were shown people with clothing in shades of blue and gray. Randomly selected group 2 subjects were shown people with clothing in shades of brown and yellow. Randomly selected group 3 subjects were shown people with clothing in shades of pink and orange. The results follow.

Group 1	Group 2	Group 3
8	7	4
7	8	9
7	7	6
7	7	7
8	5	9
8	8	8
6	5	5
8	8	8
8	7	7
7	6	5
7	6	4
8	6	5
8	6	4

1. Use the Tukey test to test all possible pairwise comparisons.
2. Are there any contradictions in the results?
3. Explain why separate *t* tests are not accepted in this situation.
4. When would Tukey's test be preferred over the Scheffé method? Explain.

See page 683 for the answers.

Exercises 12–2

1. What two tests can be used to compare two means when the null hypothesis is rejected using the one-way ANOVA *F* test?

2. Explain the difference between the two tests used to compare two means when the null hypothesis is rejected using the one-way ANOVA *F* test.

For Exercises 3 through 8, the null hypothesis was rejected. Use the Scheffé test when sample sizes are unequal or the Tukey test when sample sizes are equal, to test the differences between the pairs of means. Assume all variables are normally distributed, samples are independent, and the population variances are equal.

3. Exercise 9 in Section 12–1.

4. Exercise 12 in Section 12–1.

5. Exercise 13 in Section 12–1.

6. Exercise 17 in Section 12–1.

7. Exercise 18 in Section 12–1.

8. Exercise 20 in Section 12–1.

For Exercises 9 through 13, do a complete one-way ANOVA. If the null hypothesis is rejected, use either the Scheffé or Tukey test to see if there is a significant difference in the pairs of means. Assume all assumptions are met.

9. Emergency Room Visits Fractures accounted for 2.4% of all U.S. emergency room visits for a total of 389,000 visits for a recent year. A random sample of weekly ER visits is recorded for three hospitals in a large metropolitan area during the summer months. At $\alpha = 0.05$, is there sufficient evidence to conclude a difference in means?

Hospital X	Hospital Y	Hospital Z
28	30	25
27	18	20
40	34	30
45	28	22
29	26	18
25	31	20

Source: World Almanac.

10. Sales for Leading Companies The sales in millions of dollars for a year of a sample of leading companies are shown. At $\alpha = 0.01$, is there a significant difference in the means?

Cereal	Chocolate Candy	Coffee
578	311	261
320	106	185
264	109	302
249	125	689
237	173	

Source: Information Resources, Inc.

11. Fiber Content of Foods The number of grams of fiber per serving for a random sample of three different kinds of foods is listed. Is there sufficient evidence at the 0.05 level of significance to conclude that there is a difference in mean fiber content among breakfast cereals, fruits, and vegetables?

Breakfast cereals	Fruits	Vegetables
3	5.5	10
4	2	1.5
6	4.4	3.5
4	1.6	2.7
10	3.8	2.5
5	4.5	6.5
6	2.8	4
8		3
5		

Source: The Doctor's Pocket Calorie, Fat, and Carbohydrate Counter.

12. Per-Pupil Expenditures The expenditures (in dollars) per pupil for states in three sections of the country are listed. Using $\alpha = 0.05$, can you conclude that there is a difference in means?

Eastern third	Middle third	Western third
4946	6149	5282
5953	7451	8605
6202	6000	6528
7243	6479	6911
6113		

Source: New York Times Almanac.

13. Weekly Unemployment Benefits The average weekly unemployment benefit for the entire United States is \$314.74. Three states are randomly selected, and a sample of weekly unemployment benefits is recorded for each. At $\alpha = 0.05$, is there sufficient evidence to conclude a difference in means? If so, perform the appropriate test to find out where the difference exists.

Florida	Pennsylvania	Maine
200	300	250
187	350	195
192	295	275
235	362	260
260	280	220
175	340	290

Source: World Almanac.

14. To compare two means after a significant F test, the Scheffé test and the Tukey test were shown. Another commonly used test is called the *Bonferroni* multiple-comparison test. This test uses the t distribution, but unlike the t test shown in Chapters 8 and 9, the critical value is adjusted to compensate for multiple comparisons. This is done by dividing the critical value by the number of pairwise comparisons.

The formula for the Bonferroni test is

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{s_W^2 \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

The degrees of freedom used is $N - k$. Use the Bonferroni test to compare the means $\overline{X}_2$ and $\overline{X}_3$ in Example 12–1.

12–3 Two-Way Analysis of Variance

OBJECTIVE 3

Use the two-way ANOVA technique to determine if there is a significant difference in the main effects or interaction.

The analysis of variance technique shown previously is called a one-way ANOVA since there is only *one independent variable.* The **two-way ANOVA** is an extension of the one-way analysis of variance; it involves *two independent variables.* The independent variables are also called **factors.**

The two-way analysis of variance is quite complicated, and many aspects of the subject should be considered when you are using a research design involving a two-way

ANOVA. For the purposes of this textbook, only a brief introduction to the subject will be given.

In doing a study that involves a two-way analysis of variance, the researcher is able to test the effects of two independent variables or factors on one *dependent variable.* In addition, the interaction effect of the two variables can be tested.

Photo by Jeff Vanuga, USDA Natural Resources Conservation Service

For example, suppose a researcher wishes to test the effects of two different types of plant food and two different types of soil on the growth of certain plants. The two independent variables are the type of plant food and the type of soil, while the dependent variable is the plant growth. Other factors, such as water, temperature, and sunlight, are held constant.

To conduct this experiment, the researcher sets up four groups of plants. See Figure 12–4. Assume that the plant food type is designated by the letters A_1 and A_2 and the soil type by the Roman numerals I and II. The groups for such a two-way ANOVA are sometimes called **treatment groups.** The four groups are

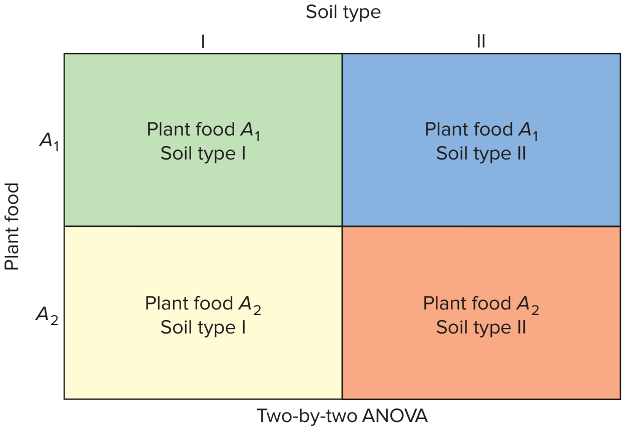
Group 1	Plant food A_1, soil type I
Group 2	Plant food A_1, soil type II
Group 3	Plant food A_2, soil type I
Group 4	Plant food A_2, soil type II

The plants are assigned to the groups at random. This design is called a 2 × 2 (read "two-by-two") design, since each variable consists of two **levels,** that is, two different treatments.

The two-way ANOVA enables the researcher to test the effects of the plant food and the soil type in a single experiment rather than in separate experiments involving the plant food alone and the soil type alone.

FIGURE 12–4

Treatment Groups for the Plant Food–Soil Type Experiment



In this case, the effect of the plant food is the change in the response variable that results from changing the level or the type of food. The effect of soil type is the change in the response variable that results from changing the level or type of soil. These two effects of the independent variable are called the **main effects.** Furthermore, the researcher can test an additional hypothesis about the effect of the *interaction* of the two variables—plant food and soil type—on plant growth. For example, is there a difference between the growth of plants using plant food A_1 and soil type II and the growth of plants using plant food A_2 and soil type I? When a difference of this type occurs, the experiment is said to have a significant **interaction effect.** The interaction effect represents the joint effect of the two factors over and above the effects of each factor considered separately. That is, the types of plant food affect the plant growth differently in different soil types. When the interaction effect is statistically significant, the researcher should not consider the effects of the individual factors without considering the interaction effect.

There are many different kinds of two-way ANOVA designs, depending on the number of levels of each variable. Figure 12–5 shows a few of these designs. As stated previously, the plant food–soil type experiment uses a 2 × 2 ANOVA.

The design in Figure 12–5(a) is called a 3 × 2 design, since the factor in the rows has three levels and the factor in the columns has two levels. Figure 12–5(b) is a 3 × 3 design, since each factor has three levels. Figure 12–5(c) is a 4 × 3 design, since the factor in the rows has four levels and the factor in the columns has three levels.

Interesting Facts

As unlikely as it sounds, lightning can travel through phone wires. You should probably hold off on taking a bath or shower as well during an electrical storm. According to the *Annals of Emergency Medicine,* lightning can also travel through water pipes.

The two-way ANOVA design has several null hypotheses. There is one for each independent variable and one for the interaction. In the plant food–soil type problem, the hypotheses are as follows:

1. The hypotheses regarding the plant food–soil type interaction effect are stated as follows.

   H_0: There is no interaction effect between type of plant food used and type of soil used on plant growth.

   H_1: There is an interaction effect between food type and soil type on plant growth.

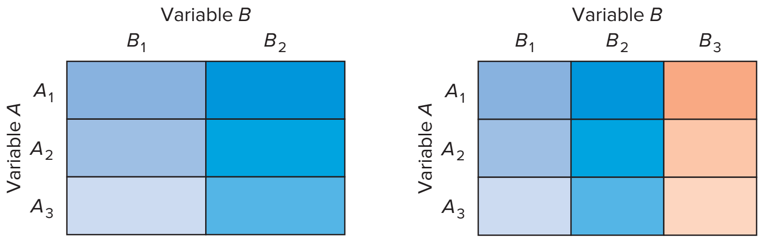
2. The hypotheses regarding plant food are stated as follows.

   H_0: There is no difference in means of heights of plants grown using different foods.

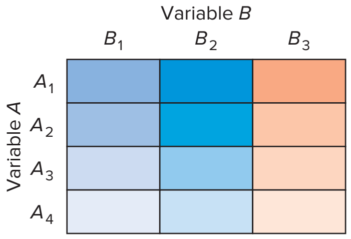
   H_1: There is a difference in means of heights of plants grown using different foods.

FIGURE 12–5

Some Types of Two-Way ANOVA Designs



(a) 3 × 2 design **(b)** 3 × 3 design



(c) 4 × 3 design

3. The hypotheses regarding soil type are stated as follows:

H_0: There is no difference in means of heights of plants grown in different soil types.

H_1: There is a difference in means of heights of plants grown in different soil types.

As with the one-way ANOVA, a between-group variance estimate is calculated, and a within-group variance estimate is calculated. An F test is then performed for each of the independent variables and the interaction. The results of the two-way ANOVA are summarized in a two-way table, as shown in Table 12–4 for the plant experiment.

TABLE 12–4 ANOVA Summary Table for Plant Food and Soil Type

Source	Sum of squares	d.f.	Mean square	*F*
Plant food				
Soil type				
Interaction				
Within (error)				
Total				

In general, the two-way ANOVA summary table is set up as shown in Table 12–5.

TABLE 12–5 ANOVA Summary Table

Source	Sum of squares	d.f.	Mean square	*F*
A	SS_A	$a-1$	MS_A	F_A
B	SS_B	$b-1$	MS_B	F_B
$A \times B$	$\mathrm{SS}_{A\times B}$	$(a-1)(b-1)$	$\mathrm{MS}_{A\times B}$	$F_{A\times B}$
Within (error)	SS_W	$ab(n-1)$	MS_W	
Total				

In the table,

SS_A = sum of squares for factor A

SS_B = sum of squares for factor B

$\mathrm{SS}_{A\times B}$ = sum of squares for interaction

SS_W = sum of squares for within-group term or error term

a = number of levels of factor A

b = number of levels of factor B

n = number of subjects in each group

$$\mathrm{MS}_A = \frac{\mathrm{SS}_A}{a-1}$$

$$\mathrm{MS}_B = \frac{\mathrm{SS}_B}{b-1}$$

$$\mathrm{MS}_{A\times B} = \frac{\mathrm{SS}_{A\times B}}{(a-1)(b-1)}$$

$$\mathrm{MS}_W = \frac{\mathrm{SS}_W}{ab(n-1)}$$

$$F_A = \frac{\mathrm{MS}_A}{\mathrm{MS}_W} \quad \text{with d.f.N.} = a-1,\ \text{d.f.D.} = ab(n-1)$$

$$F_B = \frac{\mathrm{MS}_B}{\mathrm{MS}_W} \quad \text{with d.f.N.} = b-1,\ \text{d.f.D.} = ab(n-1)$$

$$F_{A\times B} = \frac{\mathrm{MS}_{A\times B}}{\mathrm{MS}_W} \quad \text{with d.f.N.} = (a-1)(b-1),\ \text{d.f.D.} = ab(n-1)$$

The assumptions for the two-way analysis of variance are basically the same as those for the one-way ANOVA, except for sample size.

Assumptions for the Two-Way ANOVA

1. The populations from which the samples were obtained must be normally or approximately normally distributed.
2. The samples must be independent.
3. The variances of the populations from which the samples were selected must be equal.
4. The groups must be equal in sample size.

In this book, the assumptions will be stated in the exercises; however, when encountering statistics in other situations, you must check to see that these assumptions have been met before proceeding.

The two-way analysis of variance follows the regular five-step hypothesis-testing procedure.

Step 1 State the hypotheses.
Step 2 Find the critical values.
Step 3 Compute the test value.
Step 4 Make the decision.
Step 5 Summarize the results.

The computational procedure for the two-way ANOVA is quite lengthy. For this reason, it will be omitted in Example 12–5, and only the two-way ANOVA summary table will be shown. The table used in Example 12–5 is similar to the one generated by most computer programs. You should be able to interpret the table and summarize the results.

Unusual Stats

Of Americans born today, one-third of the women will reach age 100, compared to only 10% of the men, according to Ronald Klatz, M.D., president of the American Academy of Anti-Aging Medicine.

EXAMPLE 12–5 Gasoline Consumption

A researcher wishes to see whether the type of gasoline used and the type of automobile driven have any effect on gasoline consumption. Two types of gasoline, regular and high-octane, will be used, and two types of automobiles, two-wheel- and all-wheel-drive, will be used in each group. There will be two automobiles in each group, for a total of eight automobiles used. Using a two-way analysis of variance, determine if there is an interactive effect, an effect due to the type of gasoline used, and an effect due to the type of vehicle driven.

The data (in miles per gallon) are shown here, and the summary table is given in Table 12–6.

	Type of automobile	
Gas	**Two-wheel-drive**	**All-wheel-drive**
Regular	26.7	28.6
	25.2	29.3
High-octane	32.3	26.1
	32.8	24.2

TABLE 12–6 ANOVA Summary Table for Example 12–5

Source	SS	d.f.	MS	F
Gasoline A	3.920			
Automobile B	9.680			
Interaction ($A \times B$)	54.080			
Within (error)	3.300			
Total	70.980			

SOLUTION

Step 1 State the hypotheses. The hypotheses for the interaction are

H_0: There is no interaction effect between type of gasoline used and type of automobile a person drives on gasoline consumption.

H_1: There is an interaction effect between type of gasoline used and type of automobile a person drives on gasoline consumption.

The hypotheses for the gasoline types are

H_0: There is no difference between the means of gasoline consumption for two types of gasoline.

H_1: There is a difference between the means of gasoline consumption for two types of gasoline.

The hypotheses for the types of automobile driven are

H_0: There is no difference between the means of gasoline consumption for two-wheel-drive and all-wheel-drive automobiles.

H_1: There is a difference between the means of gasoline consumption for two-wheel-drive and all-wheel-drive automobiles.

Step 2 Find the critical values for each F test. In this case, each independent variable, or factor, has two levels. Hence, a 2 × 2 ANOVA table is used. Factor A is designated as the gasoline type. It has two levels, regular and high-octane; therefore, $a = 2$. Factor B is designated as the automobile type. It also has two levels; therefore, $b = 2$. The degrees of freedom for each factor are as follows:

$$\text{Factor } A\text{:} \quad \text{d.f.N.} = a - 1 = 2 - 1 = 1$$

$$\text{Factor } B\text{:} \quad \text{d.f.N.} = b - 1 = 2 - 1 = 1$$

$$\text{Interaction } (A \times B)\text{:} \quad \text{d.f.N.} = (a - 1)(b - 1) = (2 - 1)(2 - 1) = 1 \cdot 1 = 1$$

$$\text{Within (error):} \quad \text{d.f.D.} = ab(n - 1) = 2 \cdot 2(2 - 1) = 4$$

where n is the number of data values in each group. In this case, $n = 2$.

The critical value for the F_A test is found by using $\alpha = 0.05$, d.f.N. = 1, and d.f.D. = 4. In this case, $F_A = 7.71$. The critical value for the F_B test is found by using $\alpha = 0.05$, d.f.N. = 1, and d.f.D. = 4; also F_B is 7.71. Finally, the critical value for the $F_{A\times B}$ test is found by using d.f.N. = 1 and d.f.D. = 4; it is also 7.71.

Note: If there are different levels of the factors, the critical values will not all be the same. For example, if factor A has three levels and factor b has four levels, and if there are two subjects in each group, then the degrees of freedom are as follows:

$$\text{d.f.N.} = a - 1 = 3 - 1 = 2 \qquad \text{factor } A$$

$$\text{d.f.N.} = b - 1 = 4 - 1 = 3 \qquad \text{factor } B$$

$$\text{d.f.N.} = (a - 1)(b - 1) = (3 - 1)(4 - 1) = 2 \cdot 3 = 6 \qquad \text{factor } A \times B$$

$$\text{d.f.D.} = ab(n - 1) = 3 \cdot 4(2 - 1) = 12 \qquad \text{within (error) factor}$$

Step 3 Complete the ANOVA summary table to get the test values. The mean squares are computed first.

$$\text{MS}_A = \frac{\text{SS}_A}{a-1} = \frac{3.920}{2-1} = 3.920$$

$$\text{MS}_B = \frac{\text{SS}_B}{b-1} = \frac{9.680}{2-1} = 9.680$$

$$\text{MS}_{A\times B} = \frac{\text{SS}_{A\times B}}{(a-1)(b-1)} = \frac{54.080}{(2-1)(2-1)} = 54.080$$

$$\text{MS}_W = \frac{\text{SS}_W}{ab(n-1)} = \frac{3.300}{4} = 0.825$$

The F values are computed next.

$$F_A = \frac{\text{MS}_A}{\text{MS}_W} = \frac{3.920}{0.825} = 4.752 \qquad \text{d.f.N.} = a-1 = 1 \qquad \text{d.f.D.} = ab(n-1) = 4$$

$$F_B = \frac{\text{MS}_B}{\text{MS}_W} = \frac{9.680}{0.825} = 11.733 \qquad \text{d.f.N.} = b-1 = 1 \qquad \text{d.f.D.} = ab(n-1) = 4$$

$$F_{A\times B} = \frac{\text{MS}_{A\times B}}{\text{MS}_W} = \frac{54.080}{0.825} = 65.552 \qquad \text{d.f.N.} = (a-1)(b-1) = 1 \qquad \text{d.f.D.} = ab(n-1) = 4$$

The completed ANOVA table is shown in Table 12–7.

TABLE 12–7 Completed ANOVA Summary Table for Example 12–5

Source	SS	d.f.	MS	F
Gasoline A	3.920	1	3.920	4.752
Automobile B	9.680	1	9.680	11.733
Interaction ($A \times B$)	54.080	1	54.080	65.552
Within (error)	3.300	4	0.825	
Total	70.980	7		

Step 4 Make the decision. Since $F_B = 11.733$ and $F_{A\times B} = 65.552$ are greater than the critical value 7.71, the null hypotheses concerning the type of automobile driven and the interaction effect should be rejected. Since the interaction effect is statistically significant, no decision should be made about the automobile type without further investigation.

Step 5 Summarize the results. Since the null hypothesis for the interaction effect was rejected, it can be concluded that the combination of type of gasoline and type of automobile does affect gasoline consumption.

Interesting Fact

Some birds can fly as high as 5 miles.

In the preceding analysis, the effect of the type of gasoline used and the effect of the type of automobile driven are called the *main effects.* If there is no significant interaction effect, the main effects can be interpreted independently. However, if there is a significant interaction effect, the main effects must be interpreted cautiously, if at all.

To interpret the results of a two-way analysis of variance, researchers suggest drawing a graph, plotting the means of each group, analyzing the graph, and interpreting the results. In Example 12–5, find the means for each group or cell by adding the data values in each cell and dividing by n. The means for each cell are shown in the chart here.

	Type of automobile	
Gas	**Two-wheel-drive**	**All-wheel-drive**
Regular	$\overline{X}=\frac{26.7+25.2}{2}=25.95$	$\overline{X}=\frac{28.6+29.3}{2}=28.95$
High-octane	$\overline{X}=\frac{32.3+32.8}{2}=32.55$	$\overline{X}=\frac{26.1+24.2}{2}=25.15$

The graph of the means for each of the variables is shown in Figure 12–6. In this graph, the lines cross each other. When such an intersection occurs and the interaction is significant, the interaction is said to be a **disordinal interaction.** When there is a disordinal interaction, you should not interpret the main effects without considering the interaction effect.

The other type of interaction that can occur is an *ordinal interaction.* Figure 12–7 shows a graph of means in which an ordinal interaction occurs between two variables.

FIGURE 12–6

Graph of the Means of the Variables in Example 12–5

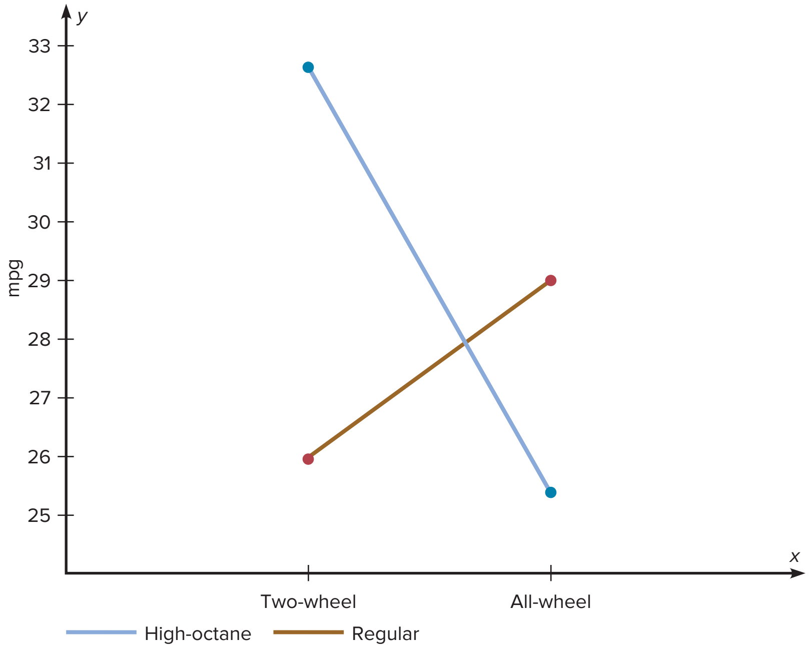


FIGURE 12–7

Graph of Two Variables Indicating an Ordinal Interaction

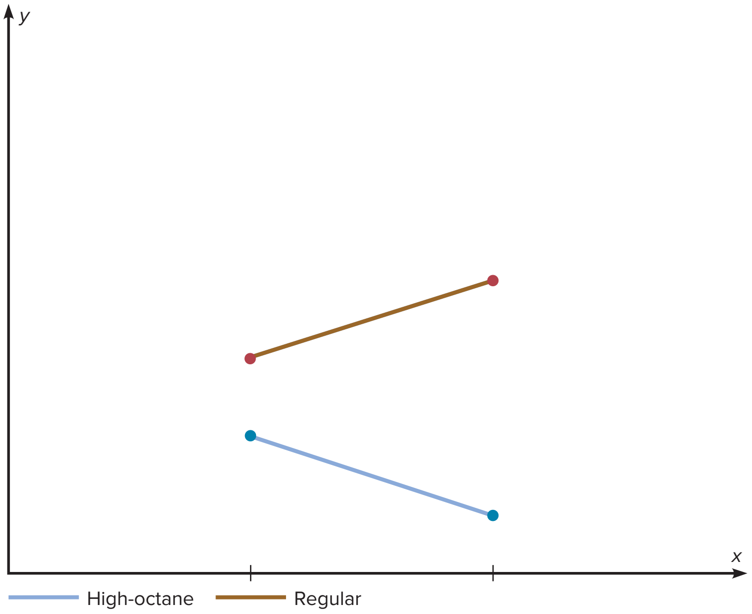
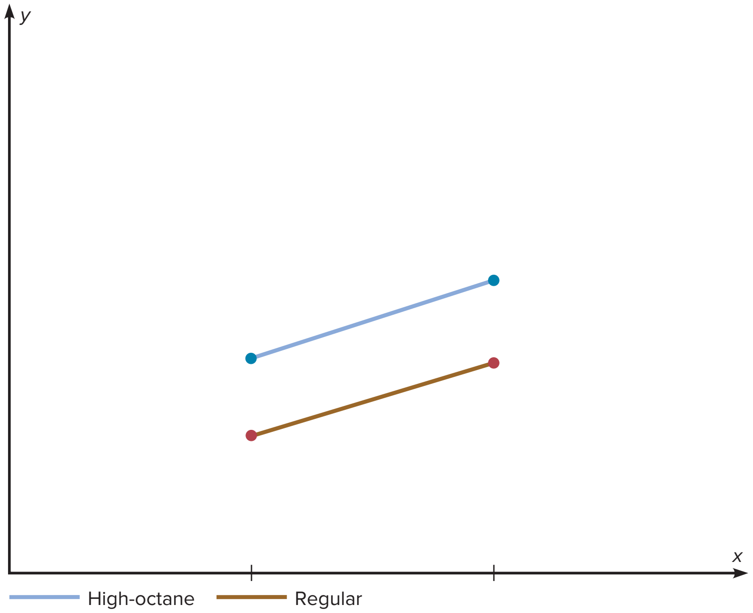


FIGURE 12–8
Graph of Two Variables Indicating No Interaction



The lines do not cross each other, nor are they parallel. If the *F* test value for the interaction is significant and the lines do not cross each other, then the interaction is said to be an **ordinal interaction** and the main effects can be interpreted independently of each other.

Finally, when there is no significant interaction effect, the lines in the graph will be parallel or approximately parallel. When this situation occurs, the main effects can be interpreted independently of each other because there is no significant interaction. Figure 12–8 shows the graph of two variables when the interaction effect is not significant; the lines are parallel.

Example 12–5 was an example of a 2 × 2 two-way analysis of variance, since each independent variable had two levels. For other types of variance problems, such as a 3 × 2 or a 4 × 3 ANOVA, interpretation of the results can be quite complicated. Procedures using tests such as the Tukey and Scheffé tests for analyzing the cell means exist and are similar to the tests shown for the one-way ANOVA, but they are beyond the scope of this textbook. Many other designs for analysis of variance are available to researchers, such as three-factor designs and repeated-measure designs; they are also beyond the scope of this book.

In summary, the two-way ANOVA is an extension of the one-way ANOVA. The former can be used to test the effects of two independent variables and a possible interaction effect on a dependent variable.

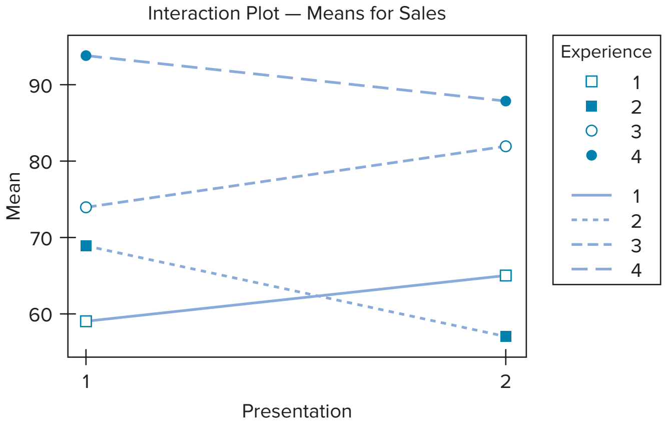
Applying the Concepts 12–3

Automobile Sales Techniques

The following outputs are from the result of an analysis of how car sales are affected by the experience of the salesperson and the type of sales technique used. Experience was broken up into four levels, and two different sales techniques were used. Analyze the results and draw conclusions about level of experience with respect to the two different sales techniques and how they affect car sales.

```
Two-Way Analysis of Variance
Analysis of Variance for Sales
Source          DF        SS         MS
Experience       3    3414.0     1138.0
Presentation     1       6.0        6.0
Interaction      3     414.0      138.0
Error           16     838.0       52.4
Total           23    4672.0
```

```
                                Individual 95% CI
Experience      Mean     ------------+------------+------------+------------+------------
1               62.0      (----------*----------)
2               63.0         (----------*----------)
3               78.0                            (-----------*-----------)
4               91.0                                           (-----------*-----------)
                         -----------+-----------+-----------+-----------+-----------
                              60.0        70.0        80.0        90.0

                                Individual 95% CI
Presentation    Mean     -----------+-----------+-----------+-----------+-----------
1               74.0           (--------------------------*--------------------------)
2               73.0     (-------------------------*-------------------------)
                         -----------+-----------+-----------+-----------+-----------
```



See page 683 for the answers.

Exercises 12–3

1. How does the two-way ANOVA differ from the one-way ANOVA?

2. Explain what is meant by *main effects* and *interaction effect.*

3. How are the values for the mean squares computed?

4. How are the *F* test values computed?

5. In a two-way ANOVA, variable *A* has three levels and variable *B* has two levels. There are five data values in each cell. Find each degrees-of-freedom value.

- *a.* d.f.N. for factor *A*
- *b.* d.f.N. for factor *B*
- *c.* d.f.N. for factor *A* × *B*
- *d.* d.f.D. for the within (error) factor

6. In a two-way ANOVA, variable *A* has six levels and variable *B* has five levels. There are seven data values in each cell. Find each degrees-of-freedom value.

- *a.* d.f.N. for factor *A*
- *b.* d.f.N. for factor *B*
- *c.* d.f.N. for factor *A* × *B*
- *d.* d.f.D. for the within (error) factor

7. What are the two types of interactions that can occur in the two-way ANOVA?

8. When can the main effects for the two-way ANOVA be interpreted independently?

For Exercises 9 through 15, perform these steps. Assume that all variables are normally or approximately normally distributed, that the samples are independent, and that the population variances are equal.

- *a.* State the hypotheses.
- *b.* Find the critical value for each *F* test.
- *c.* Complete the summary table and find the test value.
- *d.* Make the decision.
- *e.* Summarize the results. (Draw a graph of the cell means if necessary.)

9. Soap Bubble Experiments Hands-on soap bubble experiments are a great way to teach mathematics. In an effort to find the best possible bubble solution, two different soap concentrations were used along with two different amounts of glycerin additive. Students were then given a flat glass plate and a straw and were asked to blow their best bubble. The diameters of the resulting bubbles (in millimeters) are listed below. Can an interaction be concluded between the soap solution and the glycerin? Is there a difference in mean length of bubble diameter with respect to the concentration of soap to water? With respect to amount of glycerin additive? Use $\alpha = 0.05$.

	+1 Unit glycerin	**+2 Units glycerin**
Soap:water 13:25	115, 113, 105, 110	98, 100, 90, 95
Soap:water 1:2	90, 102, 100, 98	99, 100, 102, 95

10. **Increasing Plant Growth** A gardening company is testing new ways to improve plant growth. Twelve plants are randomly selected and exposed to a combination of two factors, a "Grow-light" in two different strengths and a plant food supplement with different mineral supplements. After a number of days, the plants are measured for growth, and the results (in inches) are put into the appropriate boxes.

	Grow-light 1	**Grow-light 2**
Plant food A	9.2, 9.4, 8.9	8.5, 9.2, 8.9
Plant food B	7.1, 7.2, 8.5	5.5, 5.8, 7.6

Can an interaction between the two factors be concluded? Is there a difference in mean growth with respect to light? With respect to plant food? Use $\alpha = 0.05$.

11. **Environmentally Friendly Air Freshener** As a new type of environmentally friendly, natural air freshener is being developed, it is tested to see whether the effects of temperature and humidity affect the length of time that the scent is effective. The numbers of days that the air freshener had a significant level of scent are listed below for two temperature and humidity levels. Can an interaction between the two factors be concluded? Is there a difference in mean length of effectiveness with respect to humidity? With respect to temperature? Use $\alpha = 0.05$.

	Temperature 1	**Temperature 2**
Humidity 1	35, 25, 26	35, 31, 37
Humidity 2	28, 22, 21	23, 19, 18

12. **Home-Building Times** A contractor wishes to see whether there is a difference in the time (in days) it takes two subcontractors to build three different types of homes. At $\alpha = 0.05$, analyze the data shown here, using a two-way ANOVA. See below for raw data.

Data for Exercise 12

	Home type		
Subcontractor	**I**	**II**	**III**
A	25, 28, 26, 30, 31	30, 32, 35, 29, 31	43, 40, 42, 49, 48
B	15, 18, 22, 21, 17	21, 27, 18, 15, 19	23, 25, 24, 17, 13

ANOVA Summary Table for Exercise 12

Source	**SS**	**d.f.**	**MS**	***F***
Subcontractor	1672.553			
Home type	444.867			
Interaction	313.267			
Within	328.800			
Total	2759.487			

13. **Durability of Paint** A pigment laboratory is testing both dry additives and solution-based additives to see their effect on the durability rating (a number from 1 to 10) of a finished paint product. The paint to be tested is divided into four equal quantities, and a different combination of the two additives is added to one-fourth of each quantity. After a prescribed number of hours, the durability rating is obtained for each of the 16 samples, and the results are recorded below in the appropriate space.

	Dry additive 1	**Dry additive 2**
Solution additive A	9, 8, 5, 6	4, 5, 8, 9
Solution additive B	7, 7, 6, 8	10, 8, 6, 7

Can an interaction be concluded between the dry and solution additives? Is there a difference in mean durability rating with respect to dry additive used? With respect to solution additive? Use $\alpha = 0.05$.

14. **Types of Outdoor Paint** Two types of outdoor paint, enamel and latex, were tested to see how long (in months) each lasted before it began to crack, flake, and peel. They were tested in four geographic locations in the United States to study the effects of climate on the paint. At $\alpha = 0.01$, analyze the data shown, using a two-way ANOVA shown below. Each group contained five test panels. See below for raw data.

Data for Exercise 14

	Geographic location			
Type of paint	**North**	**East**	**South**	**West**
Enamel	60, 53, 58, 62, 57	54, 63, 62, 71, 76	80, 82, 62, 88, 71	62, 76, 55, 48, 61
Latex	36, 41, 54, 65, 53	62, 61, 77, 53, 64	68, 72, 71, 82, 86	63, 65, 72, 71, 63

ANOVA Summary Table for Exercise 14

Source	SS	d.f.	MS	*F*
Paint type	12.1			
Location	2501.0			
Interaction	268.1			
Within	2326.8			
Total	5108.0			

15. Age and Sales A company sells three items: swimming pools, spas, and saunas. The owner decides to see whether the age of the sales representative and the type of item affect monthly sales. At $\alpha = 0.05$, analyze the data shown, using a two-way ANOVA. Sales are given in hundreds of dollars for a randomly selected month, and five salespeople were selected for each group.

ANOVA Summary Table for Exercise 15

Source	SS	d.f.	MS	*F*
Age	168.033			
Product	1,762.067			
Interaction	7,955.267			
Within	2,574.000			
Total	12,459.367			

Data for Exercise 15

Age of salesperson	Product		
	Pool	**Spa**	**Sauna**
Over 30	56, 23, 52, 28, 35	43, 25, 16, 27, 32	47, 43, 52, 61, 74
30 or under	16, 14, 18, 27, 31	58, 62, 68, 72, 83	15, 14, 22, 16, 27

Technology Step by Step

TI-84 Plus
Step by Step

The TI-84 Plus does not have a built-in function for two-way analysis of variance. However, the downloadable program named TWOWAY is available on the Online Learning Center. Follow the instructions for downloading the program.

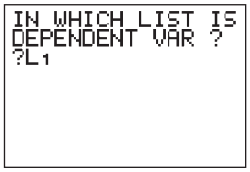
Performing a Two-Way Analysis of Variance

1. Enter the data values of the dependent variable into L_1 and the coded values for the levels of the factors into L_2 and L_3.
2. Press **PRGM,** move the cursor to the program named TWOWAY, and press **ENTER** twice.
3. Type $\mathbf{L_1}$ for the list that contains the dependent variable and press **ENTER.**
4. Type $\mathbf{L_2}$ for the list that contains the coded values for the first factor and press **ENTER.**
5. Type $\mathbf{L_3}$ for the list that contains the coded values for the second factor and press **ENTER.**
6. The program will show the statistics for the first factor.
7. Press **ENTER** to see the statistics for the second factor.
8. Press **ENTER** to see the statistics for the interaction.
9. Press **ENTER** to see the statistics for the error.
10. Press **ENTER** to clear the screen.

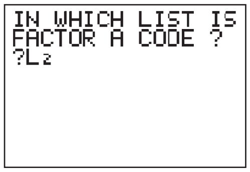
Example TI12–2

Perform a two-way analysis of variance for the gasoline data (Example 12–5 in the text). The gas mileages are the data values for the dependent variable. Factor *A* is the type of gasoline (1 for regular, 2 for high-octane). Factor *B* is the type of automobile (1 for two-wheel-drive, 2 for all-wheel-drive).

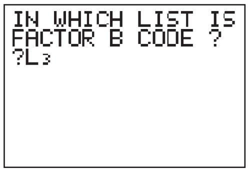
Gas mileages (L_1)	Type of gasoline (L_2)	Type of automobile (L_3)
26.7	1	1
25.2	1	1
32.3	2	1
32.8	2	1
28.6	1	2
29.3	1	2
26.1	2	2
24.2	2	2



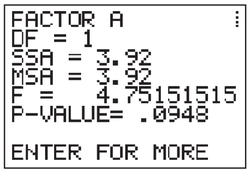
```
IN WHICH LIST IS
DEPENDENT VAR ?
?L1
```



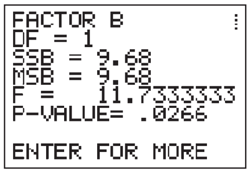
```
IN WHICH LIST IS
FACTOR A CODE ?
?L2
```



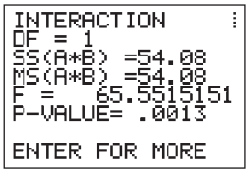
```
IN WHICH LIST IS
FACTOR B CODE ?
?L3
```



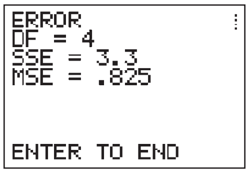
```
FACTOR A
DF = 1
SSA = 3.92
MSA = 3.92
F =   4.75151515
P-VALUE= .0948

ENTER FOR MORE
```



```
FACTOR B
DF = 1
SSB = 9.68
MSB = 9.68
F =   11.7333333
P-VALUE= .0266

ENTER FOR MORE
```



```
INTERACTION
DF = 1
SS(A*B) =54.08
MS(A*B) =54.08
F =   65.5515151
P-VALUE= .0013

ENTER FOR MORE
```



```
ERROR
DF = 4
SSE = 3.3
MSE = .825


ENTER TO END
```

EXCEL

Step by Step

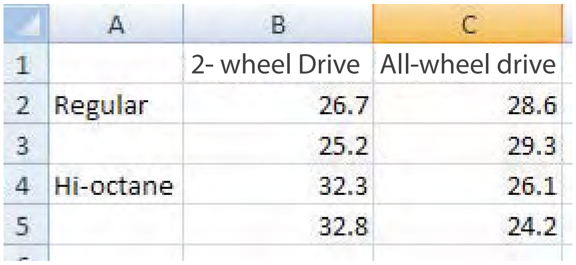
Two-Way Analysis of Variance (ANOVA)

This example pertains to Example 12–5 from the text.

Example XL12–3

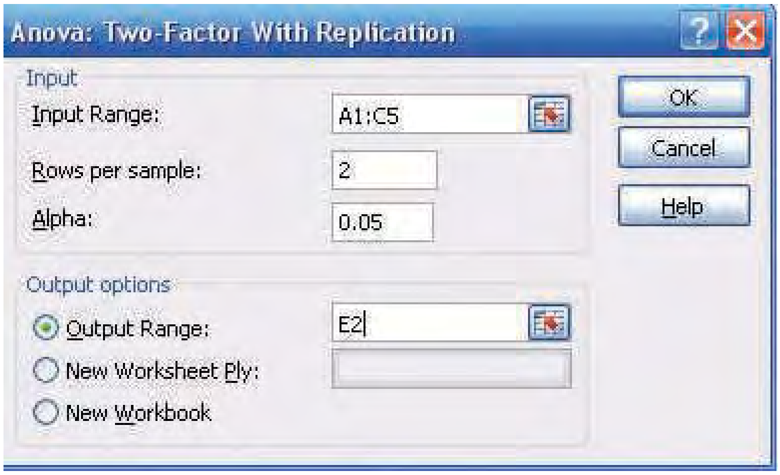
A researcher wishes to see if type of gasoline used and type of automobile driven have any effect on gasoline consumption. Use $\alpha = 0.05$.

1. Enter the data exactly as shown in the figure below in an Excel worksheet.



	A	B	C
1		2- wheel Drive	All-wheel drive
2	Regular	26.7	28.6
3		25.2	29.3
4	Hi-octane	32.3	26.1
5		32.8	24.2

2. From the toolbar, select **Data**, then **Data Analysis.**
3. Select **Anova: Two-Factor With Replication** under **Analysis tools,** then [OK].
4. In the **Anova: Single Factor dialog box,** type **A1:C5** for the **Input Range.**
5. Type **2** for the Rows per sample.
6. Type **0.05** for the **Alpha level.**
7. Under **Output options,** check **Output Range** and type **E2.**
8. Click [OK].



Anova: Two-Factor With Replication

Input

Input Range: A1:C5

Rows per sample: 2

Alpha: 0.05

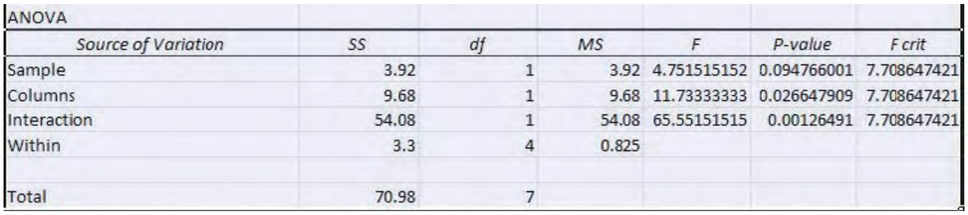
Output options

● Output Range: E2

○ New Worksheet Ply:

○ New Workbook

The two-way ANOVA table is shown below.



ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Sample	3.92	1	3.92	4.751515152	0.094766001	7.708647421
Columns	9.68	1	9.68	11.73333333	0.026647909	7.708647421
Interaction	54.08	1	54.08	65.55151515	0.00126491	7.708647421
Within	3.3	4	0.825			
Total	70.98	7				

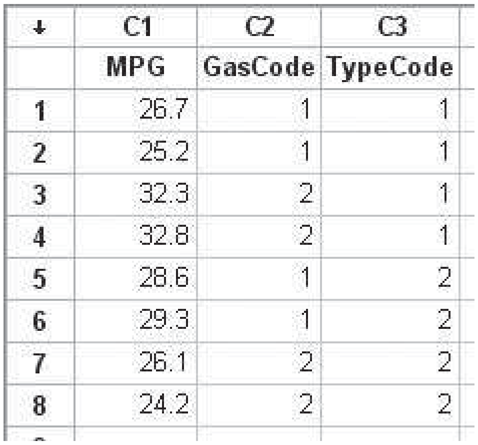
MINITAB

Step by Step

Two-Way Analysis of Variance

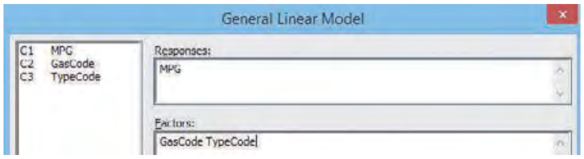
For Example 12–5, how do gasoline type and vehicle type affect gasoline mileage?

1. Enter the data into three columns of a worksheet. The data for this analysis have to be "stacked" as shown.
   a) All the gas mileage data are entered in a single column named MPG.
   b) The second column contains codes identifying the gasoline type, a 1 for regular or a 2 for high-octane.
   c) The third column will contain codes identifying the type of automobile, 1 for two-wheel-drive or 2 for all-wheel-drive.



↓	C1	C2	C3
	MPG	GasCode	TypeCode
1	26.7	1	1
2	25.2	1	1
3	32.3	2	1
4	32.8	2	1
5	28.6	1	2
6	29.3	1	2
7	26.1	2	2
8	24.2	2	2

2. Select **Stat>ANOVA>General Linear Model>Fit General Linear Model.**
   a) Double-click MPG in the list box.
   b) Double-click GasCode and TypeCode as factors.
   c) Click the Model button.
   d) Drag the mouse over GasCode and TypeCode in the upper left, then click the Add button.
   e) Click [OK].
   f) Click the Results button.
   g) Check the boxes for Analysis of variance, Model summary, and Means. Deselect any other checked boxes.
   h) Click [OK] twice.



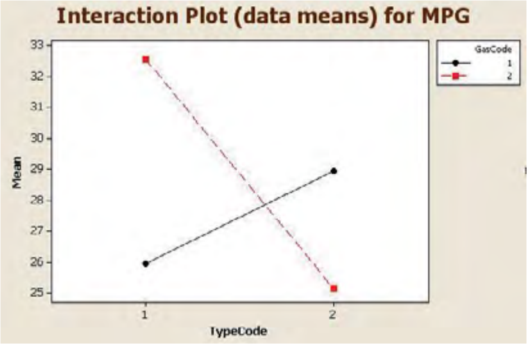
Responses: MPG

Factors: GasCode TypeCode

The session window will contain the results.

General Linear Model: MPG versus GasCode, TypeCode

```
Analysis of Variance
Source                DF     Adj SS     Adj MS   F-Value   P-Value
  GasCode              1      3.920     3.9200      4.75     0.095
  TypeCode             1      9.680     9.6800     11.73     0.027
  GasCode*TypeCode     1     54.080    54.0800     65.55     0.001
Error                  4      3.300     0.8250
Total                  7     70.980
Model Summary
       S             R-sq   R-sq(adj)   R-sq(pred)
0.908295           95.35%      91.86%       81.40%
Means               Fitted
Term                  Mean    SE Mean
GasCode
   1                27.450      0.454
   2                28.850      0.454
TypeCode
   1                29.250      0.454
   2                27.050      0.454
```



Plot Interactions

3. Select **Stat>ANOVA>Interactions Plot.**
   a) Double-click MPG for the response variable and GasCodes and TypeCodes for the factors.
   b) Click [OK].

Intersecting lines indicate a significant interaction of the two independent variables.

Summary

- The *F* test, as shown in Chapter 9, can be used to compare two sample variances to determine whether they are equal. It can also be used to compare three or more means. When three or more means are compared, the technique is called analysis of variance (ANOVA). The ANOVA technique uses two estimates of the population variance. The between-group variance is the variance of the sample means; the within-group variance is the overall variance of all the values. When there is no significant difference among the means, the two estimates will be approximately equal and the *F* test value will be close to 1. If there is a significant difference among the means, the between-group variance estimate will be larger than the within-group variance estimate and a significant test value will result. (12–1)
- If there is a significant difference among means, the researcher may wish to see where this difference lies. Several statistical tests can be used to compare the sample means after the ANOVA technique has been done. The most common are the Scheffé test and the Tukey test. When the sample sizes are the same, the Tukey test can be used. The Scheffé test is more general and can be used when the sample sizes are equal or not equal. (12–2)
- When there is one independent variable, the analysis of variance is called a one-way ANOVA. When there are two independent variables, the analysis of variance is called a two-way ANOVA. The two-way ANOVA enables the researcher to test the effects of two independent variables and a possible interaction effect on one dependent variable. If an interaction effect is found to be statistically significant, the researcher must investigate further to find out if the main effects can be examined. (12–3)

Important Terms

analysis of variance (ANOVA) 646
ANOVA summary table 648
between-group variance 647
disordinal interaction 669
factors 662
interaction effect 664
level 663
main effects 664
mean square 648
one-way ANOVA 646
ordinal interaction 670
Scheffé test 658
sum of squares between groups 648
sum of squares within groups 648
treatment groups 663
Tukey test 659
two-way ANOVA 662
within-group variance 647

Important Formulas

Formulas for the ANOVA test:

$$\overline{X}_{\text{GM}} = \frac{\Sigma X}{N}$$

$$F = \frac{s_B^2}{s_W^2}$$

where

$$s_B^2 = \frac{\Sigma n_i(\overline{X}_i - \overline{X}_{\text{GM}})^2}{k - 1} \qquad s_W^2 = \frac{\Sigma(n_i - 1)s_i^2}{\Sigma\,(n_i - 1)}$$

$$\text{d.f.N.} = k - 1 \qquad N = n_1 + n_2 + \cdots + n_k$$

$$\text{d.f.D.} = N - k \qquad k = \text{number of groups}$$

Formulas for the Scheffé test:

$$F_s = \frac{(\overline{X}_i - \overline{X}_j)^2}{s_W^2[(1/n_i) + (1/n_i)]} \quad \text{and} \quad F' = k - 1(\text{C.V.})$$

Formula for the Tukey test:

$$q = \frac{\overline{X}_i - \overline{X}_j}{\sqrt{s_W^2/n}}$$

$$\text{d.f.N.} = k \quad \text{and} \quad \text{d.f.D.} = \text{degrees of freedom for } s_W^2$$

Formulas for the two-way ANOVA:

$$\text{MS}_A = \frac{\text{SS}_A}{a - 1} \qquad F_A = \frac{\text{MS}_A}{\text{MS}_W} \qquad \begin{array}{l}\text{d.f.N.} = a - 1\\ \text{d.f.D.} = ab(n - 1)\end{array}$$

$$\text{MS}_B = \frac{\text{SS}_B}{b - 1} \qquad F_B = \frac{\text{MS}_B}{\text{MS}_W} \qquad \begin{array}{l}\text{d.f.N.} = b - 1\\ \text{d.f.D.} = ab(n - 1)\end{array}$$

$$\text{MS}_{A\times B} = \frac{\text{SS}_{A\times B}}{(a - 1)(b - 1)} \qquad F_{A\times B} = \frac{\text{MS}_{A\times B}}{\text{MS}_W} \qquad \begin{array}{l}\text{d.f.N.} = (a - 1)(b - 1)\\ \text{d.f.D.} = ab(n - 1)\end{array}$$

$$\text{MS}_W = \frac{\text{SS}_W}{ab(n - 1)}$$

Review Exercises

If the null hypothesis is rejected in Exercises 1 through 8, use the Scheffé test when the sample sizes are unequal to test the differences between the means, and use the Tukey test when the sample sizes are equal. For these exercises, perform these steps. Assume the assumptions have been met.

a. State the hypotheses and identify the claim.

b. Find the critical value(s).

c. Compute the test value.

d. Make the decision.

e. Summarize the results.

Sections 12–1 and 12–2

Use the traditional method of hypothesis testing unless otherwise specified.

1. **Lengths of Various Types of Bridges** The data represent the lengths in feet of three types of bridges in the United States. At $\alpha = 0.01$, test the claim that there is no significant difference in the means of the lengths of the types of bridges.

Simple truss	Segmented concrete	Continuous plate
745	820	630
716	750	573
700	790	525
650	674	510
647	660	480
625	640	460
608	636	451
598	620	450
550	520	450
545	450	425
534	392	420
528	370	360

Source: World Almanac and Book of Facts.

2. **Number of State Parks** The numbers of state parks found in selected states in three different regions of

the country are listed. At $\alpha = 0.05$, can it be concluded that the average number of state parks differs by region?

South	West	New England
51	28	94
64	44	72
35	24	14
24	31	52
47	40	

Source: Time Almanac.

3. **Carbohydrates in Cereals** The number of carbohydrates per serving in randomly selected cereals from three manufacturers is shown. At the 0.05 level of significance, is there sufficient evidence to conclude a difference in the average number of carbohydrates?

Manufacturer 1	Manufacturer 2	Manufacturer 3
25	23	24
26	44	39
24	24	28
26	24	25
26	36	23
41	27	32
26	25	
43		

Source: The Doctor's Pocket Calorie, Fat, and Carbohydrate Counter.

4. **Grams of Fat per Serving of Pizza** The number of grams of fat per serving for three different kinds of pizza from several manufacturers is listed. At the 0.01 level of significance, is there sufficient evidence that a difference exists in mean fat content?

Cheese	Pepperoni	Supreme/Deluxe
18	20	16
11	17	27
19	15	17
20	18	17
16	23	12
21	23	27
16	21	20

Source: The Doctor's Pocket Calorie, Fat, and Carbohydrate Counter.

5. **Iron Content of Foods and Drinks** The iron content in three different types of food is shown. At the 0.10 level of significance, is there sufficient evidence to conclude that a difference in mean iron content exists for meats and fish, breakfast cereals, and nutritional high-protein drinks?

Meats and fish	Breakfast cereals	Nutritional drinks
3.4	8	3.6
2.5	2	3.6
5.5	1.5	4.5
5.3	3.8	5.5
2.5	3.8	2.7
1.3	6.8	3.6
2.7	1.5	6.3
	4.5	

Source: The Doctor's Pocket Calorie, Fat, and Carbohydrate Counter.

6. **Temperatures in January** The average January high temperatures (in degrees Fahrenheit) for selected tourist cities on different continents are listed. Is there sufficient evidence to conclude a difference in mean temperatures for the three areas? Use the 0.05 level of significance.

Europe	Central and South America	Asia
41	87	89
38	75	35
36	66	83
56	84	67
50	75	48

Source: Time Almanac.

7. **School Incidents Involving Police Calls** A researcher wishes to see if there is a difference in the average number of times local police were called in school incidents. Random samples of school districts were selected, and the numbers of incidents for a specific year were reported. At $\alpha = 0.05$, is there a difference in the means? If so, suggest a reason for the difference.

County A	County B	County C	County D
13	16	15	11
11	33	12	31
2	12	19	3
	2	2	
	2		

Source: U.S. Department of Education.

8. **CO_2 Emissions** The CO_2 emissions in metric tons for selected countries are shown for selected years. At $\alpha = 0.10$, is there a difference in the means for the selected years?

1980	1990	2000	2010
291.2	578.6	991.0	564.6
131.7	202.1	321.5	468.7
613.6	470.6	290.5	451.0
235.0	302.2	560.3	414.5
198.8	415.4	266.2	388.7

Source: U.S. Department of Energy.

Section 12–3

9. **Review Preparation for Statistics** A statistics instructor wanted to see if student participation in review preparation methods led to higher examination scores. Five students were randomly selected and placed in each test group for a three-week unit on statistical inference. Everyone took the same examination at the end of the unit, and the resulting scores are shown. Is there sufficient evidence at $\alpha = 0.05$ to conclude an interaction between the two factors? Is there sufficient evidence to conclude a difference in mean scores based on formula delivery system? Is there sufficient evidence to conclude a difference in mean scores based on the review organization technique?

	Formulas provided	Student-made formula cards
Student-led review	89, 76, 80, 90, 75	94, 86, 80, 79, 82
Instructor-led review	75, 80, 68, 65, 79	88, 78, 85, 65, 72

10. **Effects of Different Types of Diets** A medical researcher wishes to test the effects of two different diets and two different exercise programs on the glucose level in a person's blood. The glucose level is measured in milligrams per deciliter (mg/dl). Three subjects are randomly assigned to each group. Analyze the data shown here, using a two-way ANOVA with $\alpha = 0.05$.

Exercise program	Diet A	Diet B
I	62, 64, 66	58, 62, 53
II	65, 68, 72	83, 85, 91

ANOVA Summary Table for Exercise 10

Source	**SS**	**d.f.**	**MS**	***F***
Exercise	816.750			
Diet	102.083			
Interaction	444.083			
Within	108.000			
Total	1470.916			

STATISTICS TODAY

Can Bringing Your Dog to Work Reduce Stress? —Revisited

In this study the researchers found no difference in the cortisol levels of the three groups. They could have used an analysis of variance test if the assumptions were met to determine this. There are also other possible statistical methods that can be used to compare means.

The researchers did find that, on the survey, the stress levels during the day fell by 11% among the people who had brought their dogs to work, but rose as much as 70% for the workers in other groups. The surveys were given four times a day during the week.

Data Analysis

The Data Bank is found in Appendix B, or on the World Wide Web by following links from www.mhhe.com/math/stat/bluman

1. From the Data Bank, select a random sample of subjects, and test the hypothesis that the mean cholesterol levels of the nonsmokers, less-than-one-pack-a-day smokers, and one-pack-plus smokers are equal. Use an ANOVA test. If the null hypothesis is rejected, conduct the Scheffé test to find where the difference is. Summarize the results.
2. Repeat Exercise 1 for the mean IQs of the various educational levels of the subjects.
3. Using the Data Bank, randomly select 12 subjects and randomly assign them to one of the four groups in the following classifications.

	Smoker	Nonsmoker
Male		
Female		

Use one of these variables—weight, cholesterol, or systolic pressure—as the dependent variable, and perform a two-way ANOVA on the data. Use a computer program to generate the ANOVA table.

Chapter Quiz

Determine whether each statement is true or false. If the statement is false, explain why.

1. In analysis of variance, the null hypothesis should be rejected only when there is a significant difference among all pairs of means.
2. The F test does not use the concept of degrees of freedom.
3. When the F test value is close to 1, the null hypothesis should be rejected.
4. The Tukey test is generally more powerful than the Scheffé test for pairwise comparisons.

Select the best answer.

5. Analysis of variance uses the _______ test.

a. z *c.* χ^2
b. t *d.* F

6. The null hypothesis in ANOVA is that all the means are _______.

a. Equal *c.* Variable
b. Unequal *d.* None of the above

7. When you conduct an F test, _______ estimates of the population variance are compared.

a. Two *c.* Any number of
b. Three *d.* No

8. If the null hypothesis is rejected in ANOVA, you can use the _______ test to see where the difference in the means is found.

a. z or t *c.* Scheffé or Tukey
b. F or χ^2 *d.* Any of the above

Complete the following statements with the best answer.

9. When three or more means are compared, you use the _______ technique.
10. If the null hypothesis is rejected in ANOVA, the _______ test should be used when sample sizes are equal.

For Exercises 11 through 17, use the traditional method of hypothesis testing unless otherwise specified. Assume the assumptions have been met.

11. **Gasoline Prices** Random samples of summer gasoline prices per gallon are listed for three different states. Is there sufficient evidence of a difference in mean prices? Use $\alpha = 0.01$.

State 1	State 2	State 3
3.20	3.68	3.70
3.25	3.50	3.65
3.18	3.70	3.75
3.15	3.65	3.72

12. **Voters in Presidential Elections** In a recent Presidential election, a random sample of the percentage of voters who voted is shown. At $\alpha = 0.05$, is there a difference in the mean percentage of voters who voted?

Northeast	Southeast	Northwest	Southwest
65.3	54.8	60.5	42.3
59.9	61.8	61.0	61.2
66.9	49.6	74.0	54.7
64.2	58.6	61.4	56.7

Source: Committee for the Study of the American Electorate.

13. **Ages of Late-Night TV Talk Show Viewers** A media researcher wanted to see if there was a difference in the ages of viewers of three late-night television talk shows. Three random samples of viewers were selected, and the ages of the viewers are shown. At $\alpha = 0.01$, is there a difference in the means of the ages of the viewers? Why is the average age of a viewer important to a television show writer?

Jimmy Kimmel	Jimmy Fallon	Steven Colbert
53	48	40
46	51	36
48	57	35
42	46	42
35	38	39

14. **Prices of Body Soap** A consumer group desired to compare the mean price for 12-ounce bottles of liquid body soap from two nationwide brands and one store brand. Four different bottles of each were randomly selected at a large discount drug store, and the prices are noted. At the 0.05 level of significance, is there sufficient evidence to conclude a difference in mean prices? If so, perform the appropriate test to find out where.

Brand X	Brand Y	Store brand
5.99	8.99	4.99
6.99	7.99	3.99
8.59	6.29	5.29
6.49	7.29	4.49

15. **Air Pollution** A lot of different factors contribute to air pollution. One particular factor, particulate matter, was measured for prominent cities of three continents. Particulate matter includes smoke, soot, dust, and liquid droplets from combustion such that the particle is less than 10 microns in diameter and thus capable of reaching deep into the respiratory system. The measurements are listed here. At the 0.05 level of significance, is there sufficient evidence to conclude a difference in means? If so, perform the appropriate test to find out where the differences in means are.

Asia	Europe	Africa
79	34	33
104	35	16
40	30	43
73	43	

Source: World Almanac.

16. Alumni Gift Solicitation Several students volunteered for an alumni phone-a-thon to solicit alumni gifts. The number of calls made by randomly selected students from each class is listed. At $\alpha = 0.05$, is there sufficient evidence to conclude a difference in means?

Freshmen	Sophomores	Juniors	Seniors
25	17	20	20
29	25	24	25
32	20	25	26
15	26	30	32
18	30	15	19
26	28	18	20
35			

17. Diets and Exercise Programs A researcher conducted a study of two different diets and two different exercise programs. Three randomly selected subjects were assigned to each group for one month. The values indicate the amount of weight each lost.

Exercise program \ Diet	A	B
I	5, 6, 4	8, 10, 15
II	3, 4, 8	12, 16, 11

Answer the following questions for the information in the printout shown.

a. What procedure is being used?

b. What are the names of the two variables?

c. How many levels does each variable contain?

d. What are the hypotheses for the study?

e. What are the F values for the hypotheses? State which are significant, using the P-values.

f. Based on the answers to part *e,* which hypotheses can be rejected?

Computer Printout for Problem 17

```
Datafile: NONAME.SST     Procedure: Two-way ANOVA

TABLE OF MEANS:
                                    DIET
                          A . . . . .           B . . . . .          Row Mean
 EX PROG I . . . . .        5.000                 11.000               8.000
            II . . . . .    5.000                 13.000               9.000
         Col Mean           5.000                 12.000
         Tot Mean           8.500

 SOURCE TABLE:
            Source      df      Sums of Squares      Mean Square       F Ratio        p-value
              DIET       1              147.000          147.000        21.000        0.00180
           EX PROG       1                3.000            3.000         0.429        0.53106
        DIET X EX P      1                3.000            3.000         0.429        0.53106
            Within       8               56.000            7.000
             Total      11              209.000
```

Critical Thinking Challenges

Adult Children of Alcoholics

Shown here are the abstract and two tables from a research study entitled "Adult Children of Alcoholics: Are They at Greater Risk for Negative Health Behaviors?" by Arlene E. Hall. Based on the abstract and the tables, answer these questions.

1. What was the purpose of the study?
2. How many groups were used in the study?
3. By what means were the data collected?
4. What was the sample size?
5. What type of sampling method was used?
6. How might the population be defined?
7. What may have been the hypothesis for the ANOVA part of the study?
8. Why was the one-way ANOVA procedure used, as opposed to another test, such as the t test?
9. What part of the ANOVA table did the conclusion "ACOAs had significantly lower wellness scores (WS) than non-ACOAs" come from?
10. What level of significance was used?

11. In the following excerpts from the article, the researcher states that

> *. . . using the Tukey-HSD procedure revealed a significant difference between ACOAs and non-ACOAs, p = 0.05, but no significant difference was found between ACOAs and Unsures or between non-ACOAs and Unsures.*

Using Tables 12–8 and 12–9 and the means, explain why the Tukey test would have enabled the researcher to draw this conclusion.

> ***Abstract*** *The purpose of the study was to examine and compare the health behaviors of adult children of alcoholics (ACOAs) and their non-ACOA peers within a university population. Subjects were 980 undergraduate students from a major university in the East. Three groups (ACOA, non-ACOA, and Unsure) were identified from subjects' responses to three direct questions regarding parental drinking behaviors. A questionnaire was used to collect data for the study. Included were questions related to demographics, parental drinking behaviors, and the College Wellness Check (WS), a health risk appraisal designed especially for college students (Dewey & Cabral, 1986). Analysis of variance procedures revealed that ACOAs had significantly lower wellness scores (WS) than non-ACOAs. Chi-square analyses of the individual variables revealed that ACOAs and non-ACOAs were significantly different on 15 of the 50 variables of the WS. A discriminant analysis procedure revealed the similarities between Unsure subjects and ACOA subjects. The results provide valuable information regarding ACOAs in a nonclinical setting and contribute to our understanding of the influences related to their health risk behaviors.*

TABLE 12–8 Means and Standard Deviations for the Wellness Scores (WS) Group by ($N = 945$)

Group	N	$\overline{X}$	S.D.
ACOAs	143	69.0	13.6
Non-ACOAs	746	73.2	14.5
Unsure	56	70.1	14.0
Total	945	212.3	42.1

TABLE 12–9 ANOVA of Group Means for the Wellness Scores (WS)

Source	d.f.	SS	MS	F
Between groups	2	2,403.5	1,201.7	5.9*
Within groups	942	193,237.4	205.1	
Total	944	195,640.9		

*$p < 0.01$

Source: Arlene E. Hall, "Adult Children of Alcoholics: Are They at Greater Risk for Negative Health Behaviors?" *Journal of Health Education* 12, no. 4, pp. 232–238.

Data Projects

Use a significance level of 0.05 for all tests.

1. **Business and Finance** Select 10 stocks at random from the Dow Jones Industrials, the NASDAQ, and the S&P 500. For each, note the gain or loss in the last quarter. Use analysis of variance to test the claim that stocks from all three groups have had equal performance.

2. **Sports and Leisure** Use total earnings data for movies that were released in the previous year. Sort them by rating (G, PG, PG13, and R). Is the mean revenue for movies the same regardless of rating?

3. **Technology** Use the data collected in data project 3 of Chapter 2 regarding song lengths. Consider only three genres. For example, use rock, alternative, and hip hop/rap. Conduct an analysis of variance to determine if the mean song lengths for the genres are the same.

4. **Health and Wellness** Select 10 cereals from each of the following categories: cereal targeted at children, cereal targeted at dieters, and cereal that fits neither of the previous categories. For each cereal note its calories per cup (this may require some computation since serving sizes vary for cereals). Use analysis of variance to test the claim that the calorie content of these different types of cereals is the same.

5. **Politics and Economics** Conduct an anonymous survey and ask the participants to identify which of the following categories describes them best: registered Republican, Democrat, or Independent, or not registered to vote. Also ask them to give their age to obtain your data. Use an analysis of variance to determine whether there is a difference in mean age between the different political designations.

6. **Your Class** Split the class into four groups, those whose favorite type of music is rock, whose favorite is country, whose favorite is rap or hip hop, and those whose favorite is another type of music. Make a list of the ages of students for each of the four groups. Use analysis of variance to test the claim that the means for all four groups are equal.

Answers to Applying the Concepts

Section 12–1 Colors That Make You Smarter

1. The ANOVA produces a test statistic of $F = 3.06$, with a P-value of 0.059. We would fail to reject the null hypothesis and find that there is not enough evidence to conclude at $\alpha = 0.05$ that the color of a person's clothing is related to people's perceptions of how intelligent the person looks.
2. The purpose of the study was to determine if the color of a person's clothing is related to people's perceptions of how intelligent the person looks.
3. We would have to perform three separate t tests, which would inflate the error rate.

Section 12–2 Colors That Make You Smarter

1. Tukey's pairwise comparisons show no significant difference in the three pairwise comparisons of the means.
2. This agrees with the nonsignificant results of the general ANOVA test conducted in Applying the Concepts 12–1.
3. The t tests should not be used since they would inflate the error rate.
4. We prefer the Tukey test over the Scheffé test when the samples are all the same size.

Section 12–3 Automobile Sales Techniques

There is no significant difference between levels 1 and 2 of experience. Level 3 and level 4 salespersons did significantly better than those at levels 1 and 2, with level 4 showing the best results, on average, than those at levels 1 and 2. Levels 3 and 4 are not significantly different in terms of experience. If interaction between experience and type of presentation is taken into consideration, the interaction plot shows a significant difference. The best combination seems to be level 4 experience with presentation style 1.

Hypothesis-Testing Summary 2*

7. Test of the significance of the correlation coefficient.

   Example: H_0: $\rho = 0$

   Use a t test:

   $$t = r\sqrt{\frac{n-2}{1-r^2}} \qquad \text{with d.f.} = n - 2$$

8. Formula for the F test for the multiple correlation coefficient.

   Example: H_0: $\rho = 0$

   $$F = \frac{R^2/k}{(1-R^2)/(n-k-1)}$$

   $$\text{d.f.N.} = n - k \qquad \text{d.f.D.} = n - k - 1$$

9. Comparison of a sample distribution with a specific population.

   Example: H_0: There is no difference between the two distributions.

   Use the chi-square goodness-of-fit test:

   $$\chi^2 = \sum\frac{(O-E)^2}{E}$$

   d.f. = no. of categories − 1

10. Comparison of the independence of two variables.

    Example: H_0: Variable A is independent of variable B.

    Use the chi-square independence test:

    $$\chi^2 = \sum\frac{(O-E)^2}{E}$$

    $$\text{d.f.} = (R-1)(C-1)$$

11. Test for homogeneity of proportions.

    Example: H_0: $p_1 = p_2 = p_3$

    Use the chi-square test:

    $$\chi^2 = \sum\frac{(O-E)^2}{E}$$

    $$\text{d.f.} = (R-1)(C-1)$$

12. Comparison of three or more sample means.

    Example: H_0: $\mu_1 = \mu_2 = \mu_3$

    Use the analysis of variance test:

    $$F = \frac{s_B^2}{s_W^2}$$

    where

    $$s_B^2 = \frac{\Sigma n_i(\overline{X}_i - \overline{X}_{\text{GM}})^2}{k-1}$$

    $$s_W^2 = \frac{\Sigma(n_i - 1)s_i^2}{\Sigma(n_i - 1)}$$

    $$\text{d.f.N.} = k - 1 \qquad N = n_1 + n_2 + \cdots + n_k$$

    $$\text{d.f.D.} = N - k \qquad k = \text{number of groups}$$

---

*This summary is a continuation of Hypothesis-Testing Summary 1, at the end of Chapter 9.

13. Test when the *F* value for the ANOVA is significant. Use the Scheffé test to find what pairs of means are significantly different.

$$F_s = \frac{(\overline{X}_i - \overline{X}_j)^2}{s_W^2[(1/n_i) + (1/n_j)]}$$

$$F' = (k - 1)(\text{C.V.})$$

Use the Tukey test to find which pairs of means are significantly different.

$$q = \frac{\overline{X}_i - \overline{X}_j}{\sqrt{s_W^2/n}} \qquad \begin{array}{l} \text{d.f.N.} = k \\ \text{d.f.D.} = \text{degrees of freedom for } s_W^2 \end{array}$$

14. Test for the two-way ANOVA.

Example:

H_0: There is no significant difference between the variables.

H_0: There is no interaction effect between the variables.

$$\text{MS}_A = \frac{\text{SS}_A}{a - 1}$$

$$\text{MS}_B = \frac{\text{SS}_B}{b - 1}$$

$$\text{MS}_{A\times B} = \frac{\text{SS}_{A\times B}}{(a - 1)(b - 1)}$$

$$\text{MS}_W = \frac{\text{SS}_W}{ab(n - 1)}$$

$$F_A = \frac{\text{MS}_A}{\text{MS}_W} \qquad \begin{array}{l} \text{d.f.N.} = a - 1 \\ \text{d.f.D.} = ab(n - 1) \end{array}$$

$$F_B = \frac{\text{MS}_B}{\text{MS}_W} \qquad \begin{array}{l} \text{d.f.N.} = (b - 1) \\ \text{d.f.D.} = ab(n - 1) \end{array}$$

$$F_{A\times B} = \frac{\text{MS}_{A\times B}}{\text{MS}_W} \qquad \begin{array}{l} \text{d.f.N.} = (a - 1)(b - 1) \\ \text{d.f.D.} = ab(n - 1) \end{array}$$